Optimal contracting in networks

Optimal Contracting in Networks*

Ali Jadbabaie[†] Ali Kakhbod[‡]

Abstract

We study optimal contracting between a firm selling a divisible good that exhibits positive externality and a group of agents in a social network. The extent of externality that each agent receives from the consumption of neighboring agents is privately held and is unknown to the firm. By explicitly characterizing the optimal multilateral contract, we demonstrate how inefficiency in an agent's trade propagates through the network and creates unequal and network-dependent downward distortion in other agents' trades. Furthermore, we describe bilateral contracts (non-linear pricing schemes) and characterize their explicit dependence on the network structure. We show that the firm will benefit from uncertainty in an agent's valuation of other agents' externality. We describe the profit gap between multilateral and bilateral contracts and analyze the consequences of the explicit dependence of the contracts on network structure. When the network is balanced in terms of homogeneity of agents' influence, network structure has no impact on the firm's profit for bilateral contracts. On the other hand, when the influences are heterogeneous with high dispersion (as in core-periphery networks) the restriction to bilateral contracts can result in profit losses that grow unbounded with the size of networks.

Keywords: Network Economics. Information rent. Asymmetric information. Multilateral contracts. Bilateral contracts.

JEL Classification: D82. D86. D42.

*First version June 2016, this version August 2019. We are grateful to our editor, associate editor and referees for their detailed and constructive comments. We also thank Daron Acemoglu, Amir Ajorlou, Dirk Bergemann, Ricardo Caballero, Matt Elliott, Kris Iyer, Matt Jackson, Andrey Malenko, Asu Ozdaglar, Hamid Sabourian, Alireza Tahbaz Salehi, Juuso Toikka, Alex Wolitzky and seminar participants at Massachusetts Institute of Technology (MIT), Harvard University, University of California, Berkeley, and University of Southern California for helpful comments. We also thank Emily Gallagher for excellent editing assistance. This work was supported by the Hand Foundation Fellowship and Army Research Office under MURI grant No. W911NF-12-1-0509.

[†]Institute for Data, Systems, and Society (IDSS), MIT, Cambridge, Mass., 02139, USA. (email: **jadbabai@mit.edu**).

[‡]Department of Economics, MIT, Cambridge, Mass., 02139, USA. (email: **akakhbod@mit.edu**).

1 Introduction

Over the past decade, there has been a flurry of research on economics of social networks and the monetization of social data. From display advertising to viral marketing, firms have been trying to use consumers' social interactions to increase their sales revenue, leading to exponential growth in revenue for online social networking platforms. While most revenue generated from social media accrues through advertising, there has been a major push to devise intelligent strategies for pricing products that exhibit network effects. Information about goods and services often spreads in networks as "word of mouth", while in other instances, the product itself has features that induces positive network effects. Examples include Nike+, the technology that tracks data from every run and connects runners around the world to online music services such as Apple Music and Spotify. In all these examples, firms can utilize a positive externality to sell more goods and services, since the choices of friends and acquaintances influence each consumer's decisions.

In many instances firms have data on the consumption of the products/services they sell, as well as the social network activity of their consumers. For example, online social-networking communities, such as Facebook, Twitter, Instagram, and Pinterest, allow firms to target users based on their social interactions. Some companies also provide services to firms based on information on aggregate network effects obtained by quantitative analyses of consumers' online behavior (e.g., Klout, Commun.it, Cloze.com and Kred). Firms expend major effort to exploit underlying network effects and forces in order to maximize profit. However, a major difficulty in using this information is the uncertainty as to how much agents *value* the externalities. An externality's value is often the private information of the agents. As an example, an agent may become aware of opinions or experiences of her friends about their usage of a specific good or service via her social interactions, but how much she cares (i.e., the extent of her attention) about their opinions is often her private information.

A monopoly firm selling a divisible good that demonstrates network effects naturally faces the following questions: How should firms incorporate knowledge of their customers' underlying social-network structure into their selling strategies when information on network effects is incomplete? Is uncertainty between agents about network effects beneficial to the firm? Does explicit knowledge of the network structure always matter in firms' profit-maximizing strategies?

To address these questions, we study optimal contracting (multilateral and bilateral) between a firm and a network of agents (consumers) with two distinguishing features: First, the positive externality that agents receive from their neighbors' consumption is captured by the underlying social structure that is often incomplete (i.e., not an all-to-all graph), with different agents in the network have varying degrees of influence.[1] Second, and more importantly, the extent of network externality to each agent is privately held, unknown to the firm and other agents. Our aim is to understand how network structure affects optimal contracts and to investigate the nature of resulting distortions and inefficiencies in principal-agent(s) problems, when the above two features are both present.

The goal of this paper is to study optimal contracting in networks when there is incomplete information about the strength of network effects. A major twist in our model, and a point of departure from the existing literature, is that the impact of aggregate network effect (externalities) to each agent is her private information which we interpret as the agent's *type*. A network's structure and the uncertainty in the strength of network effects are essential features of our analysis, and lead to the following implications.

1.1 Overview of the results

We start by explicitly specifying optimal multilateral contracts. Our first—rather intuitive—result shows that because of the positive network externality effect, inefficiency in an agent's trade (consumption) propagates throughout the network, causing a downward network-dependent distortion in other agents' trades. Due to the connectivity in the symmetrized network,[2] inefficiency in only one agent's trade is sufficient to cause a downward distortion in *all* agents' trades—even the trades of efficient agents.[3],[4]

[1]This heterogeneity of influence is important, as the extent of externality is often determined by the influential agents in the network. These agents are key individuals whose consumption behavior has a large impact on the consumption levels of others. Developing new methods and algorithms to identify influential/key agents from big data sets has recently attracted a lot of attention, see Probst et al. [2013] for a recent survey of this literature.

[2]The symmetrized network is defined as the sum of the network adjacency and its transpose. As we will see later, since the influences are asymmetric, the network adjacency matrix is not (necessarily) a symmetric matrix.

[3]Following Laffont and Martimort [2002], we call an agent *efficient* only when she reports the maximum type.

[4]Furthermore, we show that the maximum distortion in the whole network is caused by influential agents who are connected to other agents with high influence.

Given the complexities associated with the implementing optimal multilateral contracts and the difficulties associated with commitment and enforcement, we specify simpler, yet suboptimal, bilateral contracts, where additional restrictions are added to contract space to ensure that the maps are diagonal.[5] We explicitly characterize optimal bilateral contracts and study how uncertainty among consumers (agents) affects the firm's profit.

We show that the firm *benefits* from increasing uncertainty among consumers. The intuition is as follows: There are two sources of uncertainty in the model. One uncertainty is among the agents themselves and the other is between the firm and the agents. The uncertainty among the firm and the agents results in the canonical adverse-selection effect that manifests itself in the difference between the first-best and the second-best solutions. However, impact of the uncertainty among the agents themselves is more profound, and is captured by the structure of the first-best solution.[6] Given this observation, we show that higher uncertainty about an agent's type increases the expected quantity consumed by *all* agents, and thus, due to the strategic complement property, this effect in turn increases the firm's expected profit.[7,8] *Does network structure always affect firms' profits?* Surprisingly, when the agents' centralities in the underlying network are similar, the resulting bilateral contract is network-independent—network structure has no effect on the firm's profit. As a result, in balanced networks, explicit knowledge of network structures has no benefit for the firm.[9]

Equipped with the explicit characterizations of bilateral and multilateral contracts, we next consider how firms' profits in these contracts, consumers' social influences and network characteristics are related to one another. Of course, a firm's profit in any multilateral contract upper-bounds the profit in the corresponding bilateral contract. But, *how is this gap related to network structures?* To answer this important question, we study how

[5]In the literature, sometimes bilateral contracting problems are referred to as non-linear pricing.

[6]In the first-best solution there is no uncertainty between the firm and the agents, thus uncertainty in the first-best solution arises from the uncertainty among the agents.

[7]Similar result holds in the multilateral contracts.

[8]Moreover, our explicit characterizations of the optimal contracts show that *even a small shock to externality* will cause a potentially large change in the first- and second-best allocations as well as in agents' influences (as measured by their centralities) in the network. This structural difference particularly manifests itself in the importance of second-hop information in the structure of the bilateral contracts.

[9]We further derive simple upper and lower bounds on the firm's profit in terms of network spectral properties. These bounds show that aggregate characteristics of networks such as the smallest and largest eigenvalues, are enough to bound the profit.

network structure affects a firm's profit gap and profit ratio for multilateral and bilateral contracts.

We establish that, independent of network structures, the profit gap monotonically increases when extra links are added and networks become denser. This is intuitive because more links amplify strategic complementarity between agents, increasing the influence of network externality on each agent's payoff, leading to an increase in the profit gap.[10] To better understand this network feature of the profit gap, we also find a graph-theoretic upper bound on this profit gap that is increasing in the largest eigenvalue of the underlying network. A lower bound is also provided, which is increasing in the maximum in-degree in the network. Together, these results provide the intuition that with more concentration in networks (i.e., increasing maximum in-degree of networks) the profit gap becomes larger.

For robustness, we further analyze the ratio of profits in multilateral and bilateral contracts and show that for large balanced networks the ratio remains constant. This intuitive result proves that in large economies with limited heterogeneity in in-degrees, firms' profits in multilateral and bilateral contracts are proportional to one another. In contrast, however, we also provide a lower bound on this profit ratio in terms of the Frobenius norm of the Bonacich centrality in the symmetrized network and relate it to the extent of concentrations in networks. The main takeaway of this result is that in most networks with the mean-preserving spread of centralities, as the standard deviation of the centralities increases, the lower bound rises, increasing the profit ratio.

Finally, the above findings beg the following question: *Can the profit gap grow unboundedly?* By focusing on particular core-periphery network structures, we show that the extent of asymmetry in agents' in-degrees can have a major impact on the profit gap. Specifically, for certain networks (core-periphery structures), the profit gap can grow unbounded as the network size increases. This result provides the following important intuition: in large networks with large influence asymmetries (e.g., stark heterogeneity in in-degrees), firms' restrictions to bilateral contracts may result in major loss of profit.[11]

[10] We also show that, due to the uncertainty in network externalities, the optimality gap becomes directly related to the centrality of nodes of an auxiliary network in which certain directed paths of length 2 have become edges.

[11] Intuitively, the converse holds as well. That is, the profit gap can go to zero when externality weights are getting sufficiently small as the network size grows.

1.2 Related literature

The goal of this paper is to study optimal contracting in networks with incomplete information in the strength of interactions. A major twist in our model, and a point of departure from the existing literature, is the uncertainty in the value of network effect, which is each agent's private information.

We consider a model with strategic complementarity where an increase in the trade of others leads a given agent's higher trade to have a relatively higher payoff compared to the same agent's lower trade. This effect has been the subject of extensive work (e.g., Farrell and Saloner [1985], Katz and Shapiro [1986]). However, in the existing literature, the network effects often correspond to *all-to-all* or to complete graphs, whereas in our setup agents interact locally only with their neighbors according to the underlying network structure.

Strategic interactions under the presence of local network effects have been analyzed as *network games* by a series of papers in the past decade. These include Ballester et al. [2006], Bramoullé and Kranton [2007], Candogan et al. [2012], Bloch and Quérou [2013], Corbo et al. [2007], Galeotti and Goyal [2009], Jackson and Zenou [2015], Fainmesser and Galeotti [2016a,b]. Following the common trend in this literature, we assume the payoff function of agents takes the form of a quadratic function. In Ballester et al. [2006], the authors explicitly characterize the Nash equilibrium of a network game when agents choose their effort simultaneously. Furthermore, the authors show that the peer effect game has a unique Nash equilibrium in which each agent's effort is proportional to her Bonacich centrality in the original network. Other authors such as Bramoullé and Kranton [2007], Corbo et al. [2007], and Bramoullé et al. [2014] study a similar game in the context of the provisioning of public goods. Candogan et al. [2012] extend this framework to include optimal pricing when the firm and the agents have perfect knowledge of the network structure. The authors show that the optimal consumption of each consumer depends on the Bonacich centrality of the network, and as a result the optimal pricing strategy may involve offering discounts to agents who have a central position in the network and are a source of extra utility for their neighbors, while others who receive this extra utility will receive a markup in the price. Bloch and Quérou [2013] study a similar pricing setting and also model network effects as a deterministic graph that is commonly known to the firm and agents. They show that the value of location-based price discrimination depends on the extent of convexity of the firm's cost function. Other authors, such as Galeotti et al. [2010] and Sundararajan [2008], consider the more realistic setting of firm's limited knowledge about the original social network's structure such as knowledge of the degree distribution, and thus derive optimal mechanisms that depend on this first-degree measure of influence of an agent. Recently, Fainmesser and Galeotti [2016a] consider a setup similar to Galeotti et al. [2010] and Sundararajan [2008], which utilizes the knowledge of the degree distribution according to which consumers are aware of their own in-degree and out-degree, and are uncertain about the level of interactions of others. The authors develop price discrimination strategies based on a firm's information about consumers' influence and consumers' susceptibility to influence, evaluate the value of information on network effects for the firm, and show that value of information about consumers' influence and/or consumers' susceptibility increases in the dispersion of the in- and out-degree distributions, as well as the average level of network effects.[12] In contrast to these works, we consider optimal contracting with local externalities and *incomplete* information on the value of network effects.[13]

Our work is also in line with the body of literature on optimal contracting in principal-agent models by Mussa and Rosen [1978], Maskin and Riley [1984], Laffont and Tirole [1990], and, in particular on devising optimal contracts with externalities by Segal [1999, 2003], Csorba [2008], Bernstein and Winter [2012].[14] In particular, Segal [1999, 2003] in his seminal papers develops a model of contracting with externalities under complete information with an all-to-all (complete) network structure and characterizes the nature of the arising distortions and inefficiencies. When externalities are positive, he shows that

[12]Fainmesser and Galeotti [2016b] study the practice of influencer marketing and its effect on market efficiency in oligopoly markets. They show that firms subsidize (charge premia) consumers whose influence is above (below) that of average influential consumers. The size of premia/discounts depend on the strength of network effects and the level of information that firms have on consumers' influence.

[13]Other network-related works include: Business cycles (Acemoglu et al. [2012]), learning (Golub and Jackson [2010], Acemoglu et al. [2011], Golub and Jackson [2012]), advertising and targeted pricing (Bimpikis et al. [2016], Bloch [2016], Leduc et al. [2017], Shin [2017] Chen et al. [2018]), dynamic pricing (Ajorlou et al. [2018]), and network games (Galeotti and Mattozzi [2011], Zhou and Chen [2015], Belhaj et al. [2016]).

[14]Bernstein and Winter [2012] setup is in complete information that studies how to subsidize in order to obtain efficient coordination in a setting. Csorba [2008] extends Segal [1999, 2003] works to include incomplete information and demonstrate that the joint presence of asymmetric information and positive externality leads to a downward distortion in the welfare-maximizing allocation for all agents. Importantly, the nature of externality in those works is such that the utility of an agent depends on the behavior of the whole set of agents; however, in our setup agents interact locally only with their neighbors, and uncertainty is in network effects.

each agent's consumption level is smaller in the resulting equilibrium allocation than in the socially efficient one. In contrast to these works, we consider optimal contracting in the presence of a general *network structure* with *incomplete information* in the aggregate network effect. We particularly consider how network structures affect multilateral and bilateral contracts.

The rest of the paper is organized as follows: Section 2 presents the general model. Section 3 characterizes the optimal multilateral contract and makes a linkage to the Bonacich centrality measure. The optimal bilateral contract is specified in Section 4, and the comparison between these contracts is in Section 5. Our conclusion closes the paper. Proofs appear in the appendix.

2 Model: optimal (multilateral) contracts

A monopoly firm sells a divisible good that may be used by n consumer agents in varying quantities. The firm's marginal cost of production is normalized to zero. Agents in this market form a social network denoted by $\mathsf{G} = (V, E)$, where the vertex set $V = \{1, \cdots, n\}$ corresponds to the set of agents, and the edge set $E \subseteq V^2$ corresponds to social relationships. The corresponding adjacency in the network is captured by a matrix denoted by G (with a slight abuse of notation G denotes the network as well as the adjacency matrix corresponding to it). The ij-th entry of G, denoted by g_{ij}, represents the *strength* of the influence of a agent j on i. We assume $g_{ij} \in [0\ 1]$ for all i, j and we set $g_{ii} = 0$ for all i. For any $i, j \in V$, when $g_{ij} > 0$, agent j induces some positive externality on agent i. We further assume the underlying undirected network, which is formed by ignoring orientation of edges, is connected. The network G is common knowledge, i.e., all g_{ij} are commonly known.

We further assume the payoff function of agents is quadratic. More specifically, each agent i's payoff is specified as follows:

$$u_i(\theta_i, x_i, x_{-i}, t_i) = \underbrace{a x_i - \frac{b}{2} x_i^2}_{\text{direct utility}} + \underbrace{\theta_i \sum_{j \neq i} g_{ij} x_i x_j}_{\substack{\text{indirect utility} \\ \text{(type-dependent network effect)}}} - \underbrace{t_i}_{\text{payment}}, \qquad (1)$$

where x_i is the amount of the good she consumes, $x_{-i} \triangleq (x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$ is the consumption of other agents excluding agent i, and t_i is the disutility charged for x_i by the firm.

In addition, the parameter $\theta_i \in [\underline{\theta}\ \bar{\theta}]$ controls the (aggregate) strength of the network effect for agent i. Each agent's type θ_i is her *private information* (i.e., neither the firm nor other agents know the type), but it is commonly known that θ_i is drawn independently from a cumulative distribution F, for all i. We assume F is continuous and has a full support, with continuous density f such that $f(\cdot) > 0$. Moreover, $\phi(\tau) \triangleq \frac{1-F(\tau)}{f(\tau)}$, the reciprocal of the hazard rate, is assumed to be non-increasing for all τ.

We can interpret the payoff function of each agent i as follows. The first two terms of (1) represent the direct utility agent i derives from her trade (consumption) x_i, independent of the trades of her neighbors. The third term represents the (*type-dependent*) network-externality effect when agent i accepts the proposed offer. Since g_{ij} for all i and j are known, this captures the situation where an agent may become aware of opinions or experiences of her friends about their usage of a specific good or service via her social interactions (which is the known g_{ij} in our formulation), but how much agent i *values/cares about* the *aggregate* network effect due to her friends' consumptions (which is formulated as $\sum_{j\neq i} g_{ij}x_i x_j$ in our formulation) is her private information (which is captured by her private type θ_i). Also, we note that g_{ij} and g_{ji} might differ from one another. When $g_{ij} > 0$, trading of i and j with the firm has a strategic complement property. That is, an increase in agent j's consumption triggers an upward shift in agent i's consumption. Finally, the last term in (1) is agent i's payment to the firm.

In this economy, the firm's objective is to devise a menu of optimal incentive quantity-price pairs $\{x_i(\cdot), t_i(\cdot)\}$ for each agent i so as to maximize her total ex-ante profit defined as:

$$\mathbf{E}_{\theta\in[\underline{\theta}\ \bar{\theta}]^n}\left[\sum_{i=1}^{n} t_i(\theta)\right]. \tag{2}$$

Notations Before proceeding further, we introduce some notation that will be used throughout. We will occasionally write "$-i$" to mean agents other than agent i, and θ to denote a type profile, i.e., $\theta \in [\underline{\theta}\ \bar{\theta}]^n$. If M is a square matrix, then M^T denotes its transpose. Unless indicated otherwise, a bold letter denotes a column vector (e.g. $\mathbf{x}$), its transpose is denoted by $\mathbf{x}^T$ and the i's-th element is denoted by $[\mathbf{x}]_i$. Finally, I denotes the identity matrix.

3 Contracting game

The timing of the contracting game is as follows:

- Period 1: Each agent i observes her private type θ_i.
- Period 2: Firm announces to each agent i the menu of quantity-price pairs $\{x_i(\cdot), t_i(\cdot)\}$.
- Period 3: Each agent i based on her private type θ_i determines her optimal announcement type that is

$$\hat{\theta}_i = \arg\max_{\tau_i \in [\underline{\theta}\ \bar{\theta}]} \mathbf{E}_{\theta_{-i}}\Big[u_i(\theta_i, x_i(\tau_i, \theta_{-i}), x_{-i}(\tau_i, \theta_{-i}), t_i(\tau_i, \theta_{-i}))\Big],$$

  subject to her participation constraint

$$\mathbf{E}_{\theta_{-i}}\Big[u_i(\theta_i, x_i(\hat{\theta}_i, \theta_{-i}), x_{-i}(\hat{\theta}_i, \theta_{-i}), t_i(\hat{\theta}_i, \theta_{-i}))\Big] \geq 0,$$

  (we have assumed the reservation utility of each agent is normalized to zero).

- Period 4: The payoff of all agents and the firm are realized.

We assume each agent i, before making her decision, observes the menus offered to all agents. This assumption essentially rules out a situation where the firm and one agent trade secretly and eliminates the case where one agent's trade is contingent on the trades of the others.[15] Appealing to the Revelation Principle, in characterizing optimal quantity-price pairs $\{x_i(\cdot), t_i(\cdot)\}_{i=1}^n$, the firm can focus on direct revelation mechanisms in which the agents announce their types, and truthful reporting constitutes a Bayes-Nash equilibrium.[16] Thus, the firm's problem is rewritten as:

$$\max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^n} \quad \mathbf{E}\left[\sum_{i=1}^n t_i(\theta)\right]$$

$$\text{subject to} \quad \text{(IC):}\ \theta_i = \arg\max_{\tau_i \in [\underline{\theta}\ \bar{\theta}]} \mathbf{E}_{\theta_{-i}}\left[u_i(\theta_i, x_i(\tau_i, \theta_{-i}), x_{-i}(\tau_i, \theta_{-i}), t_i(\tau_i, \theta_{-i}))\right]\ \forall i,\ \theta_i, \tag{3}$$

$$\text{(PC):}\ \mathbf{E}_{\theta_{-i}}\left[u_i(\theta_i, x_i(\theta_i, \theta_{-i}), x_{-i}(\theta_i, \theta_{-i}), t_i(\theta_i, \theta_{-i}))\right] \geq 0\ \forall i,\ \theta_i, \tag{4}$$

[15] This is common in contracting literature, see Segal [1999, 2003].

[16] By the Revelation Principle, without loss of generality, a firm can restrict her search to design contract profiles wherein each agent i always finds it optimal to report her type θ_i truthfully.

where equations (3) and (4) are the corresponding incentive-compatibility (IC) and participation constraints (PC), respectively. As in standard adverse-selection problems, the solution of the above program in the absence of IC constraints is called the *first-best* solution, denoted by $\{x_i^{FB}(\cdot), t_i^{FB}(\cdot)\}_{i=1}^n$. And when IC constraints are present, the solution is called the *second-best* solution and is denoted by $\{x_i^{SB}(\cdot), t_i^{SB}(\cdot)\}_{i=1}^n$.

3.1 First-best and second-best solutions

Since without the IC constraints in the firm's problem, all PC constraints must bind, the first-best trade profile $\{x_i^{FB}(\cdot)\}_{i=1}^n$ is indeed (ex ante) total surplus maximizing:

$$\{x_i^{FB}(\cdot)\}_{i=1}^n \in \arg\max_{\{x_i(\cdot)\}_{i=1}^n} \mathbf{E}\sum_{i=1}^n \left[ax_i - \frac{b}{2}x_i^2 + \theta_i x_i \sum_{j\neq i} g_{ij}x_j\right]. \tag{5}$$

$\{x_i^{FB}(\cdot)\}_{i=1}^n$ serves as a benchmark. Any discrepancy between the second-best trade profile $\{x_i^{SB}(\cdot)\}_{i=1}^n$ and the benchmark solution, for each realization of θ, is called *distortion*, and is the source of inefficiency. To ensure the objective function in Eq. (5) is indeed concave and has an interior solution for each type profile $\theta \in [\underline{\theta}\ \bar{\theta}]^n$, we make the following assumption. This condition stipulates that local complementarities must be small enough compared to own concavity, which then prevents excessive feedback that can lead to the absence of a finite trade profile.

Assumption 1. *For each $i \in V$, $b > \bar{\theta}\sum_{j\neq i}(g_{ij} + g_{ji})$.*

Before characterizing the second-best solution we have the following assumption, to ensure $x_i^{SB}(\theta)$ is an interior solution[17] to the firm's problem, for all $i \in V$ and $\theta \in [\underline{\theta}\ \bar{\theta}]^n$.

Assumption 2. $\psi(\underline{\theta}) \triangleq \underline{\theta} - \frac{1-F(\underline{\theta})}{f(\underline{\theta})} \geq 0.$[18]

The term $\psi(\theta_i)$ is usually referred as a *virtual type*. The next proposition characterizes the first-best and the second-best trade profiles in terms of network structure.

[17] Without Assumption 2 we may need to do bunching.

[18] As an example, one can easily show that the following class of (β−parametric) cumulative distribution

$$F_\beta(\tau) = 1 - \left(\frac{\bar{\theta}-\tau}{\bar{\theta}-\underline{\theta}}\right)^{\frac{1}{\beta}}, \quad \beta > 0, \quad \tau \in [\underline{\theta}\ \bar{\theta}],$$

has an increasing hazard rate and, in addition, for any $\beta \leq \frac{\underline{\theta}}{\bar{\theta}-\underline{\theta}}$, Assumption 2 is satisfied.

Proposition 1. *For any type profile* $\theta \in [\underline{\theta}\ \bar{\theta}]^n$*:*

(i) the first-best trade profile is given by

$$\mathbf{x}^{FB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta)]^{-1}\mathbf{1}, \tag{6}$$

where $\mathsf{M}_\theta \triangleq diag(\theta_1, \theta_2, \cdots, \theta_n)$.

(ii) the second-best trade profile is given by

$$\mathbf{x}^{SB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)]^{-1}\mathbf{1}, \tag{7}$$

where $\mathsf{M}_\psi \triangleq diag(\psi(\theta_1), \psi(\theta_2), \cdots, \psi(\theta_n))$.

Proof. See the Appendix. □

The consequences of this explicit characterization of the first-best and the second-best trade profiles in terms of the underlying network structure are analyzed in the following sections.

3.2 Distortion vs. network structure: Role of influential agents

Given the result in Proposition 1, we now introduce a distortion vector as a function of the underlying network structure. By characterizing the distortion in terms of Bonacich centrality of agents, we determine how inefficiency in one agent's trade propagates throughout the network. We further show that the impact of this distortion propagation on the firm's profit varies depending on the agents' locations in the network, and thus we identify influential agents whose inefficiencies result in maximum losses in the firm's profit.

Let us first provide the definitions of Bonacich centrality and the distortion vector, for any type profile $\theta \in [\underline{\theta}\ \bar{\theta}]^n$.

Definition 1 (Bonacich centrality)**.** *Given a network with adjacency matrix* G*, a scalar* α*, and a vector* $\mathbf{v}$ *(weighted) Bonacich centrality with parameter* α *is defined by*[19]

$$K(\mathsf{G}, \alpha, \mathbf{v}) = (\mathsf{I} - \alpha \mathsf{G})^{-1}\mathbf{v} = [k_1\ k_2\ \cdots k_n]^T,$$

and k_i *denotes agent i's Bonacich centrality.*

[19]It is assumed that $(\mathsf{I} - \alpha\mathsf{G})^{-1}$ is well defined and nonnegative.

Definition 2. *For any* $\theta \in [\underline{\theta}\ \bar{\theta}]^n$*, the difference between the first-best and second-best is referred to as distortion and is denoted by* $\mathbf{d}(\theta)$ *as follows:*

$$\mathbf{d}(\theta) \triangleq \mathbf{x}^{FB}(\theta) - \mathbf{x}^{SB}(\theta). \tag{8}$$

Next, using the result in Proposition 1 for the first-best and second-best trade profiles and the above definitions, the distortion vector is characterized by the following Lemma.

Lemma 1. *For any type profile* $\theta \in [\underline{\theta}\ \bar{\theta}]^n$*:*

(i) $\mathbf{d}(\theta) = K\left(\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta, \frac{1}{b}, \frac{a}{b}\mathbf{1}\right) - K\left(\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi, \frac{1}{b}, \frac{a}{b}\mathbf{1}\right)$.

(ii) Distortion is downward, *i.e.,* $\mathbf{d}(\theta) \geq \mathbf{0}$.

(iii) Let $[\mathbf{d}(\theta)]_i$ *denote the distortion in agent i's trade with regard to the type profile* θ*. Then:*

$$[\mathbf{d}(\theta)]_i > 0$$

if there exists at least one agent j whose $\theta_j \neq \bar{\theta}$*.*

Proof. See the Appendix. □

We summarize the key insights of the above lemma as follows. First, because of the positive network externality in the payoff functions (see (1)), distortion is always downward (i.e. $\mathbf{d}(\theta) \geq \mathbf{0}$). Second, because of connectivity (i.e., existence of a path between any pair of agents in the symmetrized network $\mathsf{G} + \mathsf{G}^T$), inefficiency in only *one* agent's trade is *sufficient* to distort other agents' allocations downward. In other words, positive network externality, together with connectivity, implies that distortion in one agent's trade propagates throughout the network and distorts even the trade of an efficient agent whose type is the highest type $\bar{\theta}$ downward. Moreover, given the explicit characterization in part (i), the way a distortion in one agent's trade propagates throughout the network depends on her position in the symmetrized network, as captured by the node's Bonacich centrality.[20] This observation immediately gives rise to the following question: Are central agents (in terms of Bonacich centrality) in $\mathsf{G} + \mathsf{G}^T$ necessarily those key agents whose distortions have the highest effect on the firm's profit? By key agents, we naturally mean

[20] To visualize better the way distortion propagates in a network, see Example B-1 in Appendix B.

those agents whose inefficiencies create maximum distortion throughout the whole network. We show that this is not necessarily the case. To proceed, we first provide the definition of total distortion in a network for any type profile θ. The total distortion is directly related to the firm's total loss caused by propagations of distortions due to inefficiencies in agents' trades. Thus, we also specify agents whose private information (and the firm's lack of access to it) can result in the largest *loss* to the firm's profit.

Definition 3. *For any type profile $\theta \in [\underline{\theta}\ \bar{\theta}]^n$, the total distortion is the sum of the discrepancies in the first- and the second-best allocations of all agents, given by:*

$$\mathcal{T}(\theta) = \mathbf{1}^T \mathbf{d}(\theta).$$

To show that central agents in $\mathsf{G} + \mathsf{G}^T$ are *not* necessarily key agents, we use the explicit characterization of the distortion vector in Lemma 1 and determine the overall distortion as well as the loss that inefficiency of any agent's type can create, when others are all at the efficient type (i.e., $\theta_j = \bar{\theta}$ for all $j \neq i$).[21]

Proposition 2. *Consider any agent $i \in \{1, \cdots, n\}$ and let $\theta_j = \bar{\theta}$, for all $j \neq i$. Let k_r denote agent r's Bonacich centrality as characterized by $(\mathsf{I} - \frac{\bar{\theta}}{b}(\mathsf{G} + \mathsf{G}^T))^{-1}\mathbf{1} = [k_1, k_2, \cdots, k_n]$. Then, the following holds when $\Delta\theta = \bar{\theta} - \underline{\theta}$ is sufficiently small:*

(i) When $\theta_i \leq \bar{\theta}$, the overall distortion created by agent i in any network G is given by:

$$\mathcal{T}_i(\theta_i, \bar{\theta}_{-i}) = \frac{2a}{b^2}\, |\phi'(\bar{\theta})|\, (\bar{\theta} - \theta_i) \left[\sum_{j=1}^{n} k_i k_j g_{ij} \right]. \tag{9}$$

(ii) Let $\theta_i \in \{\underline{\theta}, \bar{\theta}\}$ with $Prob\{\theta_i = \underline{\theta}\} = \mathsf{v} > \frac{\Delta\theta}{\bar{\theta}}$. Then, the lack of the firm's certain knowledge (incomplete information) about agent i's type maximizes expected loss in the firm's profit when $i \in V^\star$ where

$$V^\star \triangleq \arg \max_{i \in \{1,2,\cdots,n\}} \mathbf{E}\left[\underbrace{\Pi^{FB}(\theta_i, \bar{\theta}_{-i}) - \Pi^{SB}(\theta_i, \bar{\theta}_{-i})}_{loss} \right] = \arg \max_{i \in \{1,2,\cdots,n\}} \sum_{j=1}^{n} k_i k_j g_{ij}, \tag{10}$$

[21] The assumption that all except one agent is efficient is for the purpose of obtaining a meaningful normalization. This simplifies the comparisons and serves as a simple benchmark to achieve our goal of showing that central agents in $\mathsf{G} + \mathsf{G}^T$ are *not* necessarily key agents.

where Π^{FB} and Π^{SB} denote the firm's profit due to the corresponding first-best and second-best contracts, respectively.

Proof. See the Appendix. □

The above proposition provides three results that are worth highlighting. First, the overall distortion created by agent i decreases in her type. This is intuitive because distortions created by agent i decrease as agent i becomes more efficient. The second point concerns identifying key agents whose inefficiencies create the maximum distortion in the entire network. According to Eq. (9), the maximum distortion in the whole network is due to an agent for whom the product of her and her neighbor's centralities are maximized in the symmetrized network. Finally, the maximum loss to the firm's profit occurs when the firm does not have access to the key agents' types.

Therefore, in order to identify the key agents in G, *not only each agent i's centrality in* $\mathsf{G}+\mathsf{G}^T$ *(i.e. k_i) but also her neighbors' centralities* (i.e., k_j through g_{ij}) have to be taken into account. As a consequence, the notion of a key agent in our model is different from other models with the linear-quadratic payoff functions (e.g., most notably Ballester et al. [2006], Candogan et al. [2012] and Bloch and Quérou [2013]). In these important works, key agents are directly central agents in either the underlying network G or the symmetrized network $\mathsf{G}+\mathsf{G}^T$. This structural difference that arises in our model is mainly due to the nature of uncertainty, which is placed within the network externality effect. In Appendix G we show that if uncertainty is situated in the *direct* utility, then maximum distortion becomes due to agents that are central in the symmetrized network $\mathsf{G}+\mathsf{G}^T$, following the previous works in the literature.

We highlight the impact of this observation in the next example, where we show that centrality of an agent in the symmetrized network $\mathsf{G}+\mathsf{G}^T$ is insufficient for the agent to be a key agent as far as trade is concerned.

Example 1 (Two-star Network.). *Consider the following Two-star network, $g_{ij} \in \{0,\kappa\}$, for*[22] *all i, j, capturing the interconnection among the agents. Agent 1 obtains externality from agent 2, but not vice versa. In this network, agents 1 and 2 are both equally central in $\mathsf{G}+\mathsf{G}^T$, i.e., $k_1 = k_2$. In addition, the peripheral nodes all have the same centralities, lower than the central nodes, i.e., $k_j = k_i < k_1 = k_2$, for all $i, j \in \{3,4,\cdots,2d+2\}$. Let $\Lambda(\theta_i) \triangleq \Big(2\phi'(\bar{\theta})(\theta_i - \bar{\theta})\Big)$, for*

[22] $\kappa > 0$ is small enough so that $(\mathsf{I}-\bar{\theta}(\mathsf{G}+\mathsf{G}^T))^{-1}$ is well-defined. Note also that $a = b = 1$.

all $\theta_i \in [\underline{\theta}\ \ \bar{\theta}]$. Using Proposition 2, we show that although agents 1 and 2 both have the same centralities, inefficiency in agent 1's report creates more *distortion (in the whole network) than does agent 2's. To see this, let $\theta_1 = \theta_2 = \gamma \in [\underline{\theta}\ \bar{\theta}]$, then:*

$$\mathcal{T}_1(\gamma, \bar{\theta}_{-1}) - \mathcal{T}_2(\gamma, \bar{\theta}_{-2}) = \Big(\kappa\Lambda(\gamma)\Big)\, k_1 k_2 = \underbrace{\Big(\kappa\Lambda(\gamma)\Big)}_{>0}\,(k_1)^2 > 0,$$

where $(\gamma, \bar{\theta}_{-2}) \triangleq (\bar{\theta}, \gamma, \bar{\theta}, \cdots, \bar{\theta})$ and $(\gamma, \bar{\theta}_{-1}) \triangleq (\gamma, \bar{\theta}, \bar{\theta}, \cdots, \bar{\theta})$. Notice that this discrepancy is strictly increasing in κ and k_1, resulting in a potentially drastic difference by increasing κ and/or k_1.

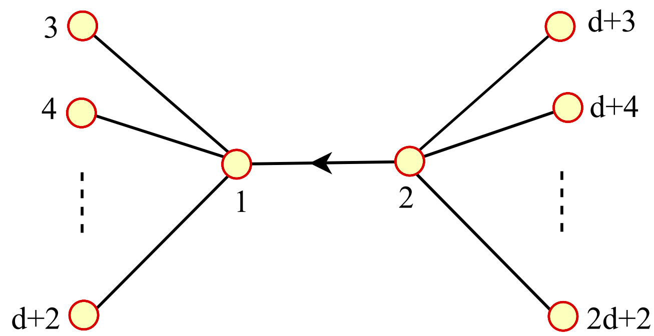

Figure 1: Two-star network. Agents 1 and 2 are both equally central in $\mathsf{G} + \mathsf{G}^T$. But distortion created in the *whole network* due to inefficiency in agent 1's type is strictly greater than agent 2's.

4 Bilateral contracts

So far, we have analyzed the optimal *multilateral* contracts in the presence of network externalities. In this section, we consider a situation where the firm is forced to offer "bilateral contracts" (which can essentially be interpreted as cases of nonlinear pricing). Although a bilateral contract is not profit-maximizing,[23] it is practically simpler to implement.[24]

[23]It is not profit maximizing, since these contracts are less constrained and there exists inefficiency even at the highest type profile.

[24]In the next section, we compare multilateral and bilateral contracts in terms of network structure.

In bilateral contracts, the firm's objective is to devise a menu of optimal incentive quantity-price pairs $\{x_i(\theta_i), t_i(\theta_i)\}$ for each agent i, where $\theta_i \in [\underline{\theta}\ \bar{\theta}]$, so as to maximize her total ex-ante profit defined as: $\sum_{i=1}^{n} \mathbf{E}[t_i(\theta_i)]$. In contrast to multilateral contracts, each agent i's menu is *only* a function of her (reported) type and *not* the type profile. Using the Revelation Principle, the firm's problem can be recast as follows:

$$\max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^{n}} \quad \sum_{i=1}^{n} \mathbf{E}[t_i(\theta_i)]$$

$$\text{subject to} \quad \text{(IC):} \quad \theta_i = \arg\max_{\tau_i \in [\underline{\theta}\ \bar{\theta}]} u_i(\theta_i, x_i(\tau_i), t_i(\tau_i)) \quad \forall i,\ \theta_i, \tag{11}$$

$$\text{(PC):} \quad u_i(\theta_i, x_i(\theta_i), t_i(\theta_i)) \geq 0 \quad \forall i,\ \theta_i, \tag{12}$$

where Eq. (11) and Eq. (12) are the corresponding incentive-compatibility and participation constraints, respectively. Moreover, the payoff function of each agent i is now updated as:

$$u_i(\theta_i, x_i, t_i) = \underbrace{a x_i - \frac{b}{2} x_i^2}_{\text{direct utility}} + \underbrace{\theta_i x_i \sum_{j \neq i} g_{ij}\ \mathbf{E}[x_j(\theta_j)]}_{\text{type-dependent indirect utility}} - \underbrace{t_i}_{\text{payment}}.$$

We preface our characterization of optimal bilateral contracts with the following remarks.

Remark 1. *Multilateral and bilateral contracts can be different, not only in the magnitude of allocations, but also in the* induced order of allocations. *While having a different allocation is natural, we note that certain network structures might also induce a different rank order on the amount of allocations (see Example B-3 in Appendix B and the formal result in Appendix F).*

Remark 2. *We note that similar to multilateral contracts, analyzing the effect of uncertainty in the full information case (the first-best) and in the imperfect-information case (the second-best) are essentially the same task. In the second-best the uncertainty is of the virtual type $\psi(\theta)$, whereas in the first-best the uncertainty is of the type θ. We analyze the impact of uncertainty on the firm's profit in Appendix D.*

Remark 3. *There is no difference (structurally) between the first- and the second-best solutions (primarily, from the network point of view). That is, the structure of the second-best is actually identical to that of the first-best, only θ changes to $\psi(\theta)$— which is due to the canonical adverse-selection effect. Consequently, the results in the imperfect-information (the second-best) case are actually the same as the results in the full-information (the first-best) case, with the minor yet important difference that uncertainty in θ (in the first-best) is replaced with the uncertainty in $\psi(\theta)$ (in the second-best). As a result, in what follows, for ease of exposition we present the results in terms of first-best solution. The corresponding second-best results with the minor change are considered in Appendix E.*

We begin with the following technical assumption. This assumption ensures that matrices that appear in the following proposition are indeed invertible and, thus, the corresponding allocations are interior and bounded solutions.[25]

Assumption 3. *Let $\rho(A)$ denote the spectral radius of A.*[26] *Then*

$$\max\{\rho\left(\frac{1}{b}\left[\mu_\psi(\mathrm{G}+\mathrm{G}^T)+\frac{\sigma_\psi^2}{b}\mathrm{G}^T\mathrm{G}\right]\right),\rho\left(\frac{1}{b}\left[\mu(\mathrm{G}+\mathrm{G}^T)+\frac{\sigma^2}{b}\mathrm{G}^T\mathrm{G}\right]\right)\}<1,$$

where

$$\mu \triangleq \mathbf{E}[\theta_i], \qquad \sigma^2 \triangleq Var(\theta_i), \qquad \mu_\psi \triangleq \mathbf{E}[\psi(\theta_i)], \qquad \sigma_\psi^2 \triangleq Var(\psi(\theta_i)).$$

The next proposition characterizes the optimal trade profile in terms of network structures.

[25] Another meaningful (but more conservative) presentation of this assumption is given by Lemma 1 in the Appendix. Assumption 3 is ensured if

$$\mu_\tau\left(\sum_{j=1}^n g_{ij}+\sum_{j=1}^n g_{ji}\right)+\frac{\sigma_\tau^2}{b}\sum_{j=1}^n\sum_{k=1}^n g_{kj}=\mu_\tau\left(d_i(in)+d_i(out)\right)+\frac{\sigma_\tau^2}{b}\sum_{i=1}^n d_i(in)<b,$$

for any i, where $d_i(in)$ is the in-degree and $d_i(out)$ is the out-degree of agent i, and τ stands for the corresponding ψ and θ (the mean and variance).

[26] Let $\{\lambda_1,\cdots,\lambda_n\}$ be the eigenvalues of a matrix $A\in\mathbf{R}^{n\times n}$. Then its spectral radius $\rho(A)$ is defined as: $\rho(A)=\max\{|\lambda_1|,\ldots,|\lambda_n|\}$.

Proposition 3. *For any i and θ_i, the optimal first trade profile is given by*

$$x_i^{FB}(\theta_i) = \frac{a}{b}(\theta_i - \mu)[\mathsf{GK}\mathbf{1}]_i + a[\mathsf{K}\mathbf{1}]_i \tag{13}$$

where $\mathsf{K} \triangleq \left[b\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1}$.[27]

Proof. See the Appendix. □

Given the above characterization, the optimal trade profile explicitly depends on the types, network structure, and the mean and variance of types. The uncertainty (i.e., the variance of the agents' types) affects the optimal trade profile through the matrix K. We emphasize that matrix K has an important role in the characterization of the optimal trade profiles in terms of the agents' locations in the underlying network.

It is also worth noting that when $[\underline{\theta}\ \bar{\theta}]$ shrinks to a single value θ, the allocation in our model becomes $x(\theta) = a[b\mathsf{I} - \theta(\mathsf{G}+\mathsf{G}^T)]^{-1}\mathbf{1}$, which is, expectedly, the same allocation as in Candogan et al. [2012] and Bloch and Quérou [2013].[28] Thus, the interesting point that the above characterization shows is that even a small uncertainty in θ (captured by σ^2) will lead to a substantial change in the structure of the allocation.[29]

Finally, given the explicit characterizations in Proposition 3, the following proposition shows that increasing uncertainty in an agent's type increases expected consumption in *all* agents. The intuition is simple. Increasing uncertainty in an agent's type increases the expected consumption of the agent.[30] Moreover, $\mathsf{G}+\mathsf{G}^T$ is a connected network. Thus, due to the strategic complementarity in consumptions, an increase in the expected consumption of the agent leads to an increase in the expected consumptions of other agents.

Proposition 4. *Given an agent i, we can state the following:*

---

[27] The optimal second-best trade profile is given by

$$x^{SB}(\theta_i) = \frac{a}{b}(\psi(\theta_i) - \mu_\psi)[\mathsf{GK}_\psi\mathbf{1}]_i + a[\mathsf{K}_\psi\mathbf{1}]_i \tag{14}$$

where $\mathsf{K}_\psi \triangleq \left[b\mathsf{I} - \mu_\psi(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1}$.

[28] We note that marginal cost of production in our model is, without loss of generality, normalized to zero.

[29] We further note that in Candogan et al. [2012] and Bloch and Quérou [2013] the authors consider *linear pricing*. We, however, in bilateral contracts, consider *non-linear pricing*.

[30] This is due to the nature of $\mathsf{K} \propto \sum_{k=0}(\mu(\mathsf{G}+\mathsf{G}^T) + \frac{\sigma^2}{b}\mathsf{G}^T\mathsf{G})^k$ and the linear-quadratic feature of the payoffs.

(i) The expected first-best consumption of agent i is given by $\mathbf{E}[x_i^{FB}(\tilde{\theta}_i)] = a[\mathsf{K}\mathbf{1}]_i$.

(ii) Increasing uncertainty in agent j's type increases the expected consumption of everyone (i.e., $\frac{\partial \mathbf{E}[x_i^{FB}(\tilde{\theta}_i)]}{\partial \sigma_j^2} \geq 0$*).*[31]

Proof. See the Appendix. □

4.1 Firm's profit

In this section we consider the firm's profit in bilateral contracts and its relation to network structures, and provide simple upper and lower bounds on the firm's profit in terms of the network's *spectral* properties.

Proposition 5. *The first-best firm's (expected) profit is given by* $\Pi_{\mathsf{G}}^{bi.} = \frac{a^2}{2}\, \mathbf{1}^T \mathsf{K}\mathbf{1}$.[32]

Proof. See the Appendix. □

Proposition 5 characterizes the firm's optimal profit with respect to network structures. Importantly, Proposition 5 shows that in the presence of uncertainty in the extent of externality *certain paths of length two*, due to the appearance of $\mathsf{G}^T\mathsf{G}$ in the matrix K (see Proposition 3), become important. The following example highlights this observation.

Example 2. *Consider an economy that consists of a firm and* 5 *consumer agents.*[33,34] *Let the interconnection among the agents be captured by a class of networks denoted by* G_τ, $\tau \in [0\ 1]$. *Let* $\mathsf{G}_\tau = \tau\mathsf{G}_1 + (1-\tau)\mathsf{G}_0$, $\forall \tau \in [0\ 1]$, *where* G_0 *and* G_1 *denote the star-outward and star-inward networks, respectively (see Fig. 2). Thus,* $\mathsf{G}_0 \rightsquigarrow \mathsf{G}_1$ *when* τ *varies from* 0 *to* 1.[35]

[31] Respectively, $\mathbf{E}[x_i^{SB}(\tilde{\theta}_i)] = a\Big[\mathsf{K}_\psi \mathbf{1}\Big]_i$. Also, a similar result holds under the second-best contract, by considering increasing uncertainty in corresponding virtual types.

[32] Consequently, using Definition 1, the firm's optimal expected profit is also given by $\Pi_{\mathsf{G}}^{bi.} = \frac{a}{2}\left\| K\left(\mu(\mathsf{G}+\mathsf{G}^T) + \frac{\sigma^2}{b}\mathsf{G}^T\mathsf{G}, \frac{1}{b}, \frac{a}{b}\mathbf{1}\right)\right\|_1$, where $\|\cdot\|_1$ denotes the L_1-norm. That is, $\|x\| = \sum_i |x_i|$, the L_1-norm is simply the sum of the absolute values. The distance derived from this norm is known as the Manhattan distance or L_1 distance.

[33] In this example we consider the firm's second-best profit. As noted before, the second- and the first-best results are structurally similar. That is, a similar result as in Proposition 5 holds in the second-best, i.e., $\Pi_{\mathsf{G}}^{bi.,SB} = \frac{a^2}{2}\, \mathbf{1}^T \mathsf{K}_\psi \mathbf{1}$ and also $\Pi_{\mathsf{G}}^{bi.,SB} = \frac{a}{2}\left\| K\left(\mathcal{G}+\mathcal{G}^T, \frac{1}{b}, \frac{a}{b}\mathbf{1}\right)\right\|_1$, where $\mathcal{G} \triangleq \mu_\psi \mathsf{G} + \frac{\sigma_\psi^2}{2b}\mathsf{G}^T\mathsf{G}$ and $\|\cdot\|_1$ denotes the L_1-norm.

[34] Number of agents is set to 5 just for ease of illustration.

[35] *Parameters*: Focusing on the network structure, we assume each agent's type is either low or high, i.e., $\theta_i \in \{\underline{\theta} = 1, \bar{\theta} = 2\}$, with probability $\text{Prob}\{\theta_i = 1\} = 1 - \text{Prob}\{\theta_i = 2\} = \mathsf{v} = \frac{3}{4}$, for all i. Thus, $\mu_\psi = \underline{\theta} = 1$ and $\sigma_\psi^2 = \left(\frac{1-\mathsf{v}}{\mathsf{v}}\right)(\Delta\theta)^2|_{\mathsf{v}=\frac{3}{4}} = \frac{1}{3}$. In addition, $b = 1, a = 10$ and $g_{ij} \in \{0, 0.1\}$, for all i, j.

Given Proposition 5, to derive the firm's optimal profit one needs to consider, for any $\tau \in [0\ 1]$

$$\mu_\psi(\mathsf{G}_\tau + \mathsf{G}_\tau^T) + \frac{\sigma_\psi^2}{b}\mathsf{G}_\tau^T\mathsf{G}_\tau|_{b=1,\mu_\psi=1,\sigma_\psi^2=\frac{1}{3}} = \mathsf{G}_\tau + \mathsf{G}_\tau^T + \frac{1}{3}\mathsf{G}_\tau^T\mathsf{G}_\tau$$
$$= \underbrace{\mathsf{G}_0 + \mathsf{G}_1}_{\textit{fixed for all } \tau} + \frac{1}{3}\mathsf{G}_\tau^T\mathsf{G}_\tau, \tag{15}$$

where the last equality follows by the construction of G_τ*. Observing this structure allows us to disentangle the effect of the second-hop neighbor's information via the* $\mathsf{G}^T\mathsf{G}$ *term. The impact on the firm's profit of changing* τ *from* 0 *to* 1 *is pictorially depicted in Fig. 2. As numerically shown in this figure, among all star networks, characterized by* G_τ*,* star-inward, *i.e.* G_1*, results in the highest (ex-ante) profit. This is due to the fact that the second-hop neighbors have the most impact when* $\tau = 1$.

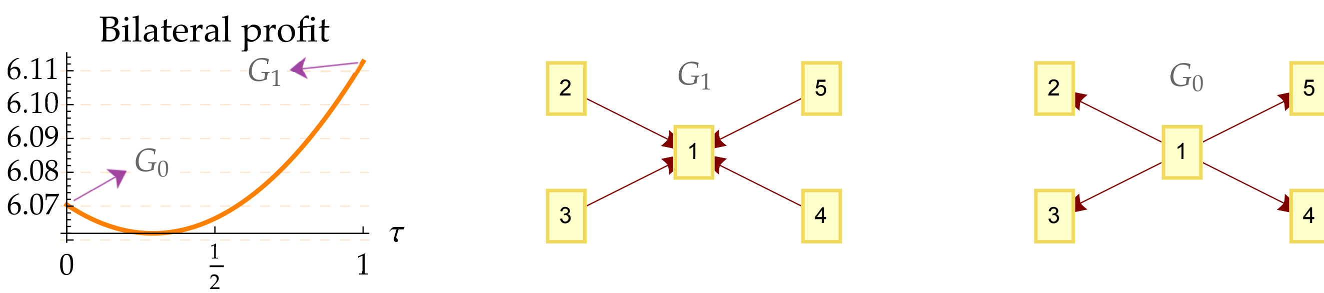

Figure 2: Interconnection among the agents.

We wrap up this section with the following proposition which characterizes quantitative upper and lower bounds on the firm's bilateral profit in terms of the network's *spectral* properties. More specifically, we discuss these bounds in terms of the maximum and minimum eigenvalues of the original symmetrized network.

In the following proposition, for clarity and ease of exposition we normalize b to 1.

Proposition 6. *Consider an* n*–agent economy with the symmetric*[36] *network structure* G*. Let*

[36] Symmetric means $\mathsf{G}^T = \mathsf{G}$.

λ_{min} and λ_{max} be the smallest and largest eigenvalues[37] *corresponding to* G*. Then for small σ^2:*

$$\frac{na^2}{2} f(\lambda_{min}) \leq \Pi^{bi.}_{\mathrm{G}} \leq \frac{na^2}{2} f(\lambda_{\max}),$$

where $f(\lambda) = \frac{1-2\mu\lambda+\sigma^2\lambda^2}{(1-2\mu\lambda)^2}$ is increasing and convex in λ and increasing in σ^2.[38]

Proof. See the Appendix. □

The above result shows that aggregate characteristics of networks, such as minimum and maximum eigenvalues, are enough to bound the profit. There is a close connection between $\lambda_{\max}$ and the average and maximum degrees d_{ave} and $d_{\max}$. Specifically, one can show that $d_{ave} \leq \lambda_{\max} \leq d_{\max}$ (followed by Prop. 2.1 in Lovász [2007]). We use this inequality to better illustrate the bounds. As shown in Proposition 6, the provided bounds are increasing and convex in λ. In particular, as the average and maximum degrees drop, the upper bound also falls, decreasing profit. Networks with low $d_{\max}$ include path-like networks, and networks with large diameters and low $d_{\max}$. While $\lambda_{\max}$ relates to maximum and average degrees, $\lambda_{\min}$ is a measure of how close the network is to being bipartite. We note that $\lambda_{\min}$ is always negative, and in bipartite networks $|\lambda_{\min}| = \lambda_{\max}$ (followed by Prop. 2.3 in Lovász [2007]). Finally, via fixing μ, with increasing σ^2 (i.e., higher uncertainty in an agent's private type), the proposed lower and upper bounds both increase, which suggests that higher uncertainty in an agent's type benefits the firm. This property is consistent with our formal result that the firm's profit is increasing in σ^2, i.e., $\frac{\partial \Pi^{bi.}_{\mathrm{G}}}{\partial \sigma^2} > 0$ (impact of uncertainty on a firm's profit is formally discussed in Appendix D).

5 Multilateral vs. bilateral contracts in networks

We now explore differences in the firm's profits between multilateral and bilateral contracts in the presence of network externality and uncertainty. In particular, we wish to address the following network-dependent questions: How and when does network structure matter? How does network structure affect the difference in the firm's profit under these various contracts? Is it possible for this difference to grow unboundedly or even go

[37] Note that since G is loop-less (i.e. $g_{ii} = 0$, for all i), thus $\sum_i \lambda_i = 0$, implying that $\lambda_{min} \leq 0 \leq \lambda_{max}$.
[38] Similar result holds for the second-best contract.

to zero? And finally, how does network structure affect the ratio of the profits? In what follows, for ease of exposition we assume, without loss of generality, that the curvature term of the utility of each agent is normalized to 1 (i.e., $b = 1$).

5.1 Balanced networks and bilateral contracts

It turns out that when the in- (or out-) degree of the symmetrized network is the same constant across agents, the network's structure has no impact on the firm's profit in the case of a bilateral contract. This means that explicit knowledge of the network structure (beyond the knowledge of uniformity of centralities) is of no value to the firm. To highlight this observation, we next define *balanced networks* to be those networks whose symmetrized graph is degree-regular.

Definition 4. *An economy with the network structure* G *is balanced if the symmetrized network* $\mathsf{G} + \mathsf{G}^T$ *has constant row sums, i.e., there exists a* $\tau > 0$ *such that*[39] $\sum_j (g_{ij} + g_{ji}) = \tau$, *for all* i.[40]

Intuitively, balanced economies are those in which no agent is overtly more influential in terms of exertion of externality.

Proposition 7. *Consider an n-agent economy with balanced structure* G. *Then, when* σ^2 *is small, the firm's profit using bilateral contract is given by* $\Pi^{bi.}_{\mathsf{G}} = \frac{a^2}{2}\left[n\zeta + \sigma^2\zeta^2 \sum_{i=1}^{n}\left(\sum_{j=1}^{n} g_{ij}\right)^2\right]$, *where* $\zeta = \frac{1}{1-\mu\tau}$.[41]

Proof. See the Appendix. □

The above proposition implies that when a firm makes use of bilateral contracts, the profit in any balanced network is only a function of the number of agents and the common degree of the symmetrized network (which is the row sum of the symmetrized network matrix). When the number of agents is fixed, all networks corresponding to the same τ result in the *same* (expected) profit. Consequently, explicit knowledge of the

[39] Note that τ is small enough so that all matrices remain invertible.

[40] We note that the term "balanced graphs" in graph theory has a slightly different meaning than the one we use here: in graph theory, a directed graph is called balanced if the in-degree and out-degree of each node is the same. Here, on the other hand, we refer to the graph as balanced if the *symmetrized network* is degree-regular graph. This means that the sum of in- and out-degrees is the same constant for all nodes in the network.

[41] A similar result holds for the second-best contract, when θ is replaced with $\psi(\theta)$.

network structure beyond the common degree is of no value to the firm. This result is particularly useful in practice, it says that when uncertainty among the agents is small and the firm makes use of bilateral contracts (i.e., non-linear pricing), then the firm's profit does not depend on the details of the network structure. This, however, may not be the case under multilateral contracts (see Fig. 3). The following corollary summarizes the above-discussed result.

Corollary 1. *All balanced networks in which each agent has an equal in-degree (e.g. cycle, wheel, regular networks) result in the same (expected) profit as when the firm uses bilateral contracts.*

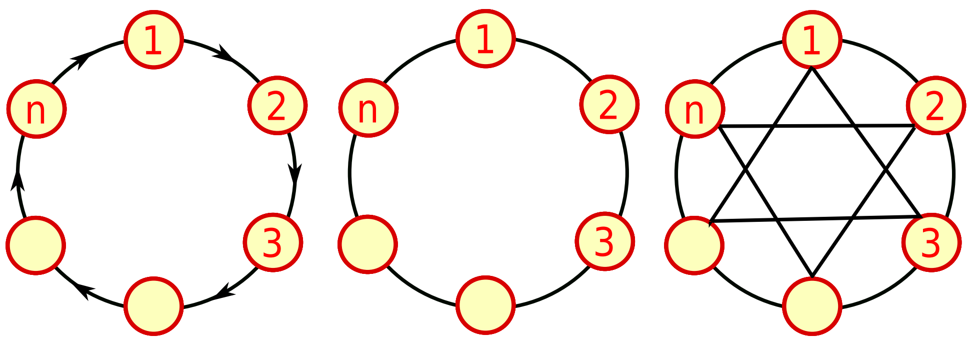

Figure 3: All of these networks with bilateral contracts result in the same profit; however, for multilateral contracts they may behave differently.

5.2 Difference in firm's profit (Profit gap): Bilateral vs. multilateral contracts

We now explore the difference in a firm's profits under multilateral and bilateral contracts as a function of network structure. In particular, the following lemma provides a closed-form expression that characterizes the suboptimality gap.

Lemma 2 (Profit gap)**.** *Suppose there exists an $\hat{m} > 0$, such that $\mathbf{E}[(\theta_i - \mu)^k] < (\hat{m}\sigma)^k$, for all i and*[42] *$k \geq 3$. For any network G and for small σ^2, let $\mathsf{S} \triangleq (I - \mu(\mathsf{G} + \mathsf{G}^T))$, and $K \triangleq$*

[42]One can easily show this constraint is satisfied by many distributions like uniform and power distributions, for appropriately chosen parameters.

$(\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T))^{-1}\mathbf{1} = \mathsf{S}^{-1}\mathbf{1}$ *be the centrality of the symmetrized graph, and define* $R_i \triangleq E_i\mathsf{G}+\mathsf{G}^T E_i$, *where* E_i *is the matrix with only the* i^{th} *diagonal set as 1 and other entries as zero, then*[43,44]

$$\Pi_{\mathsf{G}}^{multi.} - \Pi_{\mathsf{G}}^{bi.} = \left(\frac{a^2}{2}\sigma^2\right) K^T \left\{ \sum_{i=1}^{n} R_i(\mathsf{S}^{-1}-\mathsf{I})R_i + diag[(\mathsf{G}\circ\mathsf{G})\mathbf{1}] \right\} K. \quad (16)$$

Proof. See the Appendix. □

With the lemma above, we can now characterize the suboptimality of bilateral contracts. Some results are immediate. First, it is immediate[45] that $\Pi_{\mathsf{G}}^{\text{multi.}} \geq \Pi_{\mathsf{G}}^{\text{bi.}}$ and also the gap increases with the extent of uncertainty, i.e., $\frac{\partial\{\Pi_{\mathsf{G}}^{\text{multi.}}-\Pi_{\mathsf{G}}^{\text{bi.}}\}}{\partial\sigma^2} > 0$. Second, this profit gap is directly related to the Bonacich centrality of the symmetrized network.

Using Lemma 2, the following proposition characterizes several comparative statics of the profit gap in terms of network structure and its spectral properties.

Proposition 8 (Profit gap and network properties)**.** *Consider an n-agent economy with the interaction matrix* G.[46]

(i) Suppose there is no link (externality) from j to i in G, *and let* $\mathsf{G}+\{g_{ij}\}$ *denote the new matrix induced by introducing the link* $\{ij\}$. *Assuming all the invertibility assumptions are preserved, then* $\Pi_{\mathsf{G}}^{multi.} - \Pi_{\mathsf{G}}^{bi.} < \Pi_{\mathsf{G}+\{g_{ij}\}}^{multi.} - \Pi_{\mathsf{G}+\{g_{ij}\}}^{bi.}$.

(ii) If $0<\alpha<1$, *then* $\Pi_{\alpha\mathsf{G}}^{multi.} - \Pi_{\alpha\mathsf{G}}^{bi.} < \alpha^2\left(\Pi_{\mathsf{G}}^{multi.} - \Pi_{\mathsf{G}}^{bi.}\right)$.

(iii) Let G *be symmetric*[47] *and* λ_{max} *denote the largest eigenvalue of* G. *Then*

$$0 \leq \Pi_{\mathsf{G}}^{multi.} - \Pi_{\mathsf{G}}^{bi.} \leq \left(\frac{na^2}{2}\sigma^2\right) h(\lambda_{\max}),$$

where $h(\lambda) \triangleq \lambda^2\,\frac{3+2\mu\lambda}{(1-2\mu\lambda)^3}$ *that is increasing and convex in* λ.

---

[43] The notation $\circ$ denotes the element-wise Hadamard product. Given matrices $A=(a_{ij})$ and $C=(c_{ij})$ of order $n\times n$, $A\circ C=(a_{ij}c_{ij})_{ij}$, where $a_{ij}c_{ij}$ is a scalar and $A\circ C$ is of order $n\times n$.

[44] A similar result holds in the second-best contract with θ replaced by $\psi(\theta)$.

[45] Notice that since S^{-1} is well defined and positive, $\mathsf{S}^{-1}-\mathsf{I}=\sum_{i\geq 1}(\mu(\mathsf{G}+\mathsf{G}^T))^k$ is a positive matrix.

[46] The following results hold under the second-best contract with θ replaced by $\psi(\theta)$.

[47] Symmetric means $\mathsf{G}^T=\mathsf{G}$.

(iv) Let $g_{ij} \in \{0, g\}$, for all i, j. Then[48]*:*

$$\Pi_{\mathrm{G}}^{multi.} - \Pi_{\mathrm{G}}^{bi.} \geq \frac{a^2}{2}\sigma^2\mu^2 \left(d_{max}^3\, g^4\right),$$

where d_{max} denotes the maximum in-degree in G.

Proof. See the Appendix. □

Intuitively, independent of network structures, the profit gap monotonically increases when extra links are added and networks become denser. This is because more links amplify strategic complementarity between agents, increasing the influence of network externality on each agent's payoff and leading to an increase in the profit gap. Similarly, the profit gap deceases with lower weights g_{ij}. In this regard, Part (ii) shows that the influence of weight reduction on the profit gap is at least quadratic. Parts (iii) and (iv) provide quantitative upper and lower bounds on the profit gap in terms of the network's *spectral* properties. The upper bound is increasing in maximum eigenvalues of underlying networks. Therefore, when the underlying network structure becomes more dispersed (e.g., path-like networks), the proposed upper bound in most networks falls, decreasing the profit gap. Part (iv) provides a lower bound in terms of the mean and variance of the types and maximum in-degrees of underlying networks. This result implies that with greater concentration in networks (i.e., increasing maximum in-degree of networks) the proposed lower bound increases, suggesting that the profit gap can be large in dense networks.

Can the profit gap explode or vanish? In the following example, we establish that when networks are heterogeneous with *high dispersion* (as in core-periphery networks), the restriction to bilateral contracts can result in profit losses that grow unboundedly with the size of the networks. The converse holds as well. That is, expectedly, when the influencing weights (i.e., $g_{i,j}$) become proportionally small with the size of the networks, the profit gap converges to zero. All of these statements are formally proved in Appendix C.

Example 3. *Let us start with a formal definition of star-family networks:*

[48]Note that g is potentially a function of n and is small enough to preserve all the invertibility assumptions.

Definition 5 (Star-family.). *Let $\mathsf{G}_0(n)$ and $\mathsf{G}_1(n)$ denote, respectively, star-outward and star-inward graphs over n nodes. The star family $\mathfrak{G}$ includes all the δ-convex combinations of the two-star structures defined as $\mathfrak{G}(n) = \{\mathsf{G}_\delta(n) : \ \exists\ \delta \in [0,1]\ s.t.\ \mathsf{G}_\delta(n) = \delta \mathsf{G}_1(n) + (1-\delta)\mathsf{G}_0(n)\}$.*[49]

The following figure depicts the profit gap *due to the use of bilateral contracts in star networks with 50, 100 and 150 nodes.*[50] *The two extremes are star outward, $\delta = 0$, and star inward, $\delta = 1$, respectively, and the rest are the convex combinations of the two extremes. As the following figure shows, maximum loss (within the star-family) occurs in star-inwards, which is formally shown in Appendix C.*[51] *Moreover, the figure highlights the fact that with an increasing number of agents, the profit gap may explode. This statement, of course, is not always true and depends on network structure. In Appendix C, we show that the explosion in the profit gap can occur in networks that have major asymmetry in their in-degrees. When there is an agent dominant in her in-degree (e.g., star-inward) then the loss in the firm's profit due to the simpler bilateral contracts may become unbounded, and, thus, the firm's restriction to the simpler bilateral contracts may result in major losses in the firm's profit.*

5.3 Profit ratio

To provide robust intuition about the impact of network structure on bilateral and multilateral contracts, in this subsection we briefly consider how heterogeneity of influence in networks affects the *ratio* of the bilateral and multilateral profits. We first consider large balanced structures and then find upper and lower bounds on the profit ratio in general networks in terms of various spectral properties of networks.

[49] Notice that the star-family of networks is centrality-preserving (i.e., $K = (I - \mu(\mathsf{G} + \mathsf{G}^T))^{-1}\mathbf{1}$ is the same for all the networks within the core-periphery (star) family). This is because all the networks in $\mathfrak{G}(n)$ have the same symmetrized adjacency (i.e., $\mathsf{G}_\delta(n) + \mathsf{G}_\delta(n) = \mathsf{G}_{\delta'}(n) + \mathsf{G}_{\delta'}(n)$ for all $\delta, \delta' \in [0,1]$).

[50] Other parameters: $b = \sqrt{n}$, $\mu = 1/2$, $\sigma^2 = 1/12$, and $a^2/2 = 1$.

[51] It is worth noting that the main driver of this result is the second term of Eq. (16) in Lemma 2. Clearly, in the class of core-periphery structures, the maximum centrality is in the core node, which we denote it by k_1, and thus by the symmetry in the periphery nodes in $\mathsf{S}^{-1}\mathbf{1} = [k_1\ k_2\ \cdots\ k_n]^T$, we must have $k_1 > k_2 = \cdots = k_n$. Next, considering the second term of Eq. (16) (in Lemma 2), the Hadamard product $\mathsf{G} \circ \mathsf{G}$ implies that

$$K^T \mathrm{diag}[(\mathsf{G} \circ \mathsf{G})\mathbf{1}]K = \sum_{i=1}^{n} k_i^2 \parallel \mathrm{deg}_{in}(i) \parallel_2^2 = \mathbf{1}^T \mathsf{S}^{-1} \mathrm{diag}\left(\sum_{j=1}^{n} g_{1j}^2, \sum_{j=1}^{n} g_{2j}^2, \cdots, \sum_{j=1}^{n} g_{nj}^2 \right) \mathsf{S}^{-1}\mathbf{1}.$$

Thus, due to $k_1 > k_2 = \cdots = k_n$, this term is maximized for star-inward, because $\sum_{j=1}^{n} g_{1j}^2$ is maximized in star-inward.

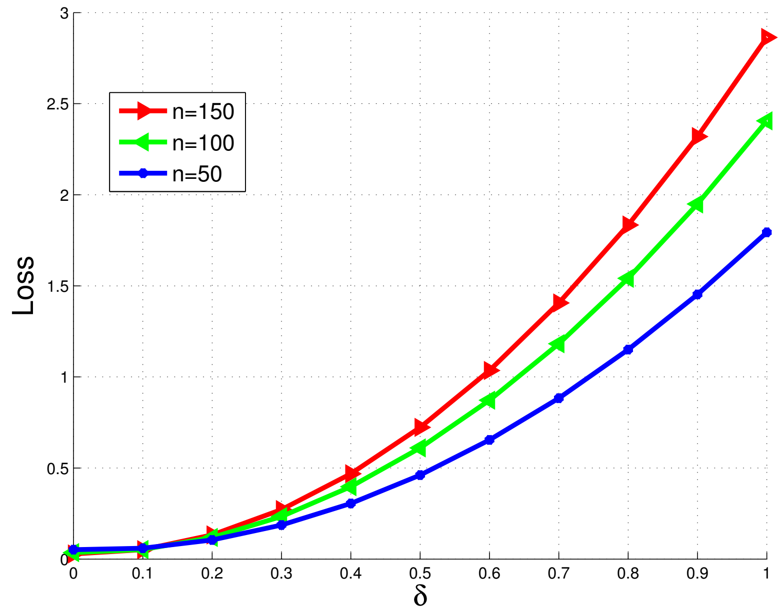

Figure 4: The plot visualizes the losses from using bilateral contracts for the family of centrality-preserving star networks. The highest profit loss happens in star (in)ward, when $\delta = 1$. Moreover, as the number of agents, i.e. n, increases, the amount of loss monotonically increases.

Proposition 9. *Consider the sequence of balanced networks* $\{\mathsf{G}(n)\}_{n=2}^{\infty}$ *with the same in-degree at each node*[52]*. As n grows:*

$$\lim_{n\to\infty} \frac{\Pi^{multi.}_{\mathsf{G}(n)}}{\Pi^{bi.}_{\mathsf{G}(n)}} = O(1).$$

Proof. See the Appendix. □

The above intuitive result demonstrates that in *large networks* with homogeneous degree of influences (i.e., balanced structures, as defined in Definition 4) the firm's profits in bilateral and multilateral contracts are of the same order (meaning that their ratio is of order 1). Hence, in large economies with *limited heterogeneity in in-degrees* these contracts are (in terms of profit) proportional to one another.[53] We next consider how heterogeneity

[52] Meaning that there exists a $0 < \kappa < 1$ so that $\sum_j g_{ij} = \kappa\tau$, for all i. We further assume $\mathsf{G}+\mathsf{G}^T$ is k regular, where k is finite.

[53] In Appendix E, we formally show the above observation is also true in cycle (clock-wise) networks, see

of influence in networks affects the profit *ratio*.

Profit ratio: upper and lower bounds. The following propositions characterize upper and lower bounds on the ratio of the firm's profits under multilateral and bilateral contracts. These bounds are in terms of the graph-spectral properties.

Proposition 10 (Upper bound). *Consider an $n-$ agent economy with the network structure* G. *Let λ_{max} denote the maximum eigenvalue of the symmetrized network* $\mathsf{G}+\mathsf{G}^T$. *Then*

$$\frac{\Pi_{\mathsf{G}}^{multi.}}{\Pi_{\mathsf{G}}^{bi.}} \leq 1+\sigma^2 g(\lambda_{\max}),$$

where $g(\lambda_{\max}) \triangleq \frac{\lambda_{\max}^2}{(1-\mu\lambda_{\max})^3}$ *is increasing in* $\lambda_{\max}$.

Proof. See the Appendix. □

The above simple upper bound is in terms of the symmetrized network's *spectral* radius. Let $d_{\max} = \max_i \sum_j (g_{ij}+g_{ji})$ denote the maximum degree in $\mathsf{G}+\mathsf{G}^T$. Since (by Proposition 2.1 in Lovász [2007]) $\max\{\sqrt{d_{\max}}, \text{average degree in } \mathsf{G}+\mathsf{G}^T\} \leq \lambda_{\max} \leq d_{\max}$, and $g(\cdot)$ is an increasing function, thus as the symmetrized network $\mathsf{G}+\mathsf{G}^T$ becomes more *dispersed* (e.g., path-like networks, networks with large diameters with low $d_{\max}$), the proposed upper bound in most networks falls, decreasing the profit ratio.

Next, we provide a lower bound on the profit ratio.

Proposition 11 (Lower bound). *Consider an $n-$ agent economy with the network structure* G. *Let* $K \triangleq (\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T))^{-1}\mathbf{1}$ *be the centrality of the symmetrized network. Then*

$$\frac{\Pi_{\mathsf{G}}^{multi.}}{\Pi_{\mathsf{G}}^{bi.}} \geq 1+\sigma^2 \frac{\|diag(K)\ \mathsf{G}\|_F^2}{\|K\|_1+\|K\|_2^2},$$

where $\|\cdot\|_F$ stands for the Frobenius norm, and $diag(K) = diag(k_1,k_2,\cdots,k_n)$. Particularly, if $\|d_i\|_2^2 = 1$, for all i, (where $d_i = (g_{i1},g_{i2},\cdots,g_{ni})$, $i-$th row of G*), then $\|diag(K)\ \mathsf{G}\|_F^2 = \|K\|_2^2$. Thus:*

$$\frac{\Pi_{\mathsf{G}}^{multi.}}{\Pi_{\mathsf{G}}^{bi.}} \geq 1+\sigma^2 \frac{\frac{\|K\|_2^2}{\|K\|_1}}{1+\frac{\|K\|_2^2}{\|K\|_1}}. \tag{17}$$

Proposition 16.

Proof. See the Appendix. □

Finally, the above lower bound in. Eq. (17) is increasing in $\|K\|_2^2$. Therefore, the above result provides the following intuition that in most networks with a mean-preserving spread of centralities (i.e., keep $\|K\|_1$ fixed), as the standard deviation of the centralities increases (i.e., increasing $\|K\|_2^2$) the lower bound rises, increasing the profit ratio (fixing $\|K\|_1$, the larger $\frac{\|K\|_2^2}{\|K\|_1}$, the more disperse the network becomes).

6 Conclusion

We study optimal contracting strategies for a firm selling a divisible good that exhibits positive externality to a finite number of consumers in a social network. A special feature of the model, and a point of departure from the existing literature, is the magnitude of network externality being the private information of each agent.

We explicitly characterize the firm's optimal multilateral and bilateral contracts as a function of the underlying network structure. Due to the presence of positive network effects, inefficiency in one agent's trade propagates throughout the network and creates an unequal and network-dependent downward distortion in *all* the agents' trades, even agents with the highest type. The distortion vector can be characterized in terms of a Bonacich centrality of the symmetrized network. In particular, inefficiencies in the trade of highly central agents who are connected to other central agents create the maximum distortion, causing the maximum loss to the firm's profits. We also explicitly characterize optimal bilateral contracts. Our explicit characterization shows that, surprisingly, uncertainty in the value of the network externality among consumers is beneficial to the firm—meaning that the firm will be better off *not* to invest in reducing uncertainty among the agents.

Focusing on the suboptimality of the bilateral contracts, we explicitly characterize the profit gap between the multilateral and bilateral contracts in terms of network structures. We show that heterogeneity in the centralities of different agents plays an important role in the profit gap. When the network structure is balanced, the firm's profit under bilateral contracts becomes independent of network structures. Furthermore, the optimality gap increases when new links are introduced. Focusing on the core-periphery structures, we also show that the optimality gap can grow unboundedly when the network size increases. Finally, we show that increasing the standard deviation of the centralities leads to an increase in the profit ratios of bilateral and multilateral contracts.

Our results shed light on scenarios in which it is profitable for firms to invest in finding the social network of their consumers and when it is not worth making such investments.

References

D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian learning in social networks. *Review of Economic Studies*, 78:1201–1236, 2011.

D. Acemoglu, V. Carvalho, A. Ozdaglar, and A Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

A. Ajorlou, A. Jadbabaie, and A. Kakhbod. Dynamic pricing in social networks: The word of mouth effect. *Management Science*, 64:971–979, 2018.

C. Ballester, A. Calvo-Armengol, and Y. Zenou. Who's who in networks. wanted: The key player. *Econometrica*, 74:1403–1417, 2006.

M. Belhaj, S. Bervoets, and F. Deroian. Efficient networks in games with local complementarities. *Theoretical Economics*, 11(1):357–380, 2016.

S. Bernstein and E. Winter. Contracting with heterogeneous externalities. *American Economic Journal: Microeconomics*, 4(2):50–76, 2012.

K. Bimpikis, A. Ozdaglar, and E. Yildiz. Competitive targeted advertising over networks. *Operation Research*, 64(3):705–720, 2016.

F. Bloch. *Targeting and pricing in social networks. In: Y. Bramoullé, A. Galeotti and B. Rogers (Eds.)*. Oxford Handbook of the Economics of Networks, Oxford: Oxford University Press, 2016.

F. Bloch and N. Quérou. Pricing in social networks. *Games and Economic Behavior*, 80: 263–281, 2013.

Y. Bramoullé and R. Kranton. Public goods in networks. *Journal of Economic Theory*, 135: 478–494, 2007.

Y. Bramoullé, R. Kranton, and M. D'Amours. Strategic interaction and networks. *American Economic Review*, 104(3):898–930, 2014.

O. Candogan, K. Bimpikis, and A. Ozdaglar. Optimal pricing in networks with externalities. *Operation Research*, 60(4):883–905, 2012.

Y. Chen, Y. Zenou, and J. Zhou. Competitive pricing strategies in social networks. *RAND Journal of Economics*, 2018.

J. Corbo, A. Calvo-Armengol, and D. Parkes. The importance of network topology in local contribution games. *Proceedings of the 3rd international Workshop on Internet and Network Economics*, 2007.

G. Csorba. Screening contracts in the presence of positive network effects. *International Journal of Industrial Organization*, 26(1):213–226, 2008.

I.P. Fainmesser and A. Galeotti. Pricing network effects. *Review of Economic Studies*, 83 (1):165–198, 2016a.

I.P. Fainmesser and A. Galeotti. Pricing network effects: Competition. *Working paper*, 2016b.

J. Farrell and G. Saloner. Standardization, compatibility, and innovation. *The RAND Journal of Economics*, 16(1):70–83, 1985.

A. Galeotti and S. Goyal. Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics*, 40(3):509–532, 2009.

A. Galeotti and A. Mattozzi. Personal influence: Social context and political competition. *American Economic Journal: Microeconomics*, 3(1):307–327, 2011.

A. Galeotti, S. Goyal, M. Jackson, F. Vega-Redondo, and L. Yariv. Network games. *Review of Economic Studies*, 77(1):218–244, 2010.

B. Golub and M.O. Jackson. Naive learning in social networks: Convergence, influence, and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2:112–149, 2010.

B. Golub and M.O. Jackson. How homophily affect the speed of learning and best response dynamics. *Quarterly Journal of Economics*, 127:1287–1338, 2012.

R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

M. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ., 2008.

M. Jackson and Y. Zenou. *Games on networks. In P. Young and S. Zamir (Eds.)*. Handbook of Game Theory, Vol. 4, Amsterdam: Elsevier Publisher, 2015.

M. Katz and C. Shapiro. Technology adoption in the presence of network externalities. *Journal of Political Economy*, 94(4):822–841, 1986.

J. Laffont and D. Martimort. *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, 2002.

J. Laffont and J. Tirole. Optimal bypass and cream skimming. *The American Economic Review*, 80:1040–1061, 1990.

M. V. Leduc, M. O. Jackson, and R. Johari. Pricing and referrals in diffusion on networks. *Games and Economic Behavior*, 104:568–594, 2017.

L. Lovász. Eigenvalues of graphs. 2007.

E. Maskin and J. Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15:171–196, 1984.

M. Mussa and C. Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18: 301–317, 1978.

F. Probst, L. Grosswiele, and R. Pfleger. Who will lead and who will follow: Identifying influential users in online social networks. *Business and Information Systems Engineering*, 5(3):179–193, 2013.

I. Segal. Contracting with externalities. *Quarterly Journal of Economics*, 114:337–388, 1999.

I. Segal. Coordination and discrimination in contracting with externalities: Divide and conquer. *Journal of Economic Theory*, 113:147–181, 2003.

E. Shin. Monopoly pricing and diffusion of social network goods. *Games and Economic Behavior*, 102:162–178, 2017.

A. Sundararajan. Local network effects and complex network structure. *The B.E. Journal of Theoretical Economics*, 7(1), 2008.

J. Zhou and Y. Chen. Key leaders in social networks. *Journal of Economic Theory*, 157: 212–235, 2015.

A Appendix

In the following lemma we present an alternative condition for Assumption 3 to hold.

Lemma 1. *In order to obtain* $\rho\left(\frac{1}{b}\left[\mu_\tau(\mathsf{G}+\mathsf{G}^T)+\frac{\sigma_\tau^2}{b}\mathsf{G}^T\mathsf{G}\right]\right)<1$, *it suffices to have for any* i:

$$\mu_\tau\left(\sum_{j=1}^{n} g_{ij}+\sum_{j=1}^{n} g_{ji}\right)+\frac{\sigma_\tau^2}{b}\sum_{j=1}^{n}\sum_{k=1}^{n} g_{kj}=\mu_\tau\left(d_i(in)+d_i(out)\right)+\frac{\sigma_\tau^2}{b}\sum_{i=1}^{n} d_i(in)<b,$$

where $d_i(in)$ *is the in-degree and* $d_i(out)$ *is the out-degree of agent* i, *and* τ *stands for the corresponding* ψ *and* θ *(the mean and variance).*

Proof. Let us denote $S \triangleq \frac{1}{b}\left[\mu_\tau(\mathsf{G}+\mathsf{G}^T)+\frac{\sigma_\tau^2}{b}\mathsf{G}^T\mathsf{G}\right]$. Let $v=[v_1\ v_2\ \cdots\ v_n]^T$ be an eigenvector of S and its corresponding eigenvalue be λ. Thus, by definition, we have

$$Sv=\lambda v.$$

Let $v_i \triangleq \max\{|v_1|,|v_2|,\cdots,|v_n|\}$, where $|A|$ denotes the absolute value of A. Hence

$$|\lambda v_i|=|S_i v|=|\sum_{j=1}^{n} s_{ij}v_j|\leq|\sum_{j=1}^{n} s_{ij}||v_i|, \tag{18}$$

where S_i denotes the i^{th} row of S. Let $\hat{g}_{ij}$ denote an entry of G^T located at i^{th} row and j^{th}

column, thus, $\hat{g}_{ij} = g_{ji}$. Therefore,

$$
\begin{aligned}
\sum_{j=1}^{n} s_{ij} &= \frac{1}{b}\left[\mu_\tau\left(\sum_{j=1}^{n} g_{ij} + \sum_{j=1}^{n} g_{ji}\right) + \frac{\sigma_\tau^2}{b}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} \hat{g}_{ik} g_{kj}\right)\right] \\
&= \frac{1}{b}\left[\mu_\tau\left(\sum_{j=1}^{n} g_{ij} + \sum_{j=1}^{n} g_{ji}\right) + \frac{\sigma_\tau^2}{b}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} g_{ki} g_{kj}\right)\right] \\
&\leq \frac{1}{b}\left[\mu_\tau\left(\sum_{j=1}^{n} g_{ij} + \sum_{j=1}^{n} g_{ji}\right) + \frac{\sigma_\tau^2}{b}\sum_{j=1}^{n}\sum_{k=1}^{n} g_{kj}\right],
\end{aligned}
$$

where the last inequality follows since $g_{ij} \in [0\ 1]$, thus $g_{ki}g_{kj} \leq g_{kj}$, $\forall\ k, j$ and i.

Now, if $\frac{1}{b}\left[\mu_\tau\left(\sum_{j=1}^{n} g_{ij} + \sum_{j=1}^{n} g_{ji}\right) + \frac{\sigma_\tau^2}{b}\sum_{j=1}^{n}\sum_{k=1}^{n} g_{kj}\right] < 1$, then the above inequality implies that $|\sum_{j=1}^{n} s_{ij}| < 1$. Therefore, due to Eq. (18), we have $|\lambda| < 1$, and consequently, $\rho(S) < 1$.

We further note that each eigenvalue of $\mathsf{I} - S$ is equal to $1 - \lambda$ where λ is an eigenvalue of S. Thus $\rho(S) < 1$ implies that all the eigenvalues of $\mathsf{I} - S$ are non-zero and, therefore, $\mathsf{I} - S$ is invertible. □

B Appendix: extra examples

Example B-1 (Path Network): Nature of distortion This example considers distortion propagations in a path network when there is uncertainty in the extent of interactions. In particular, it shows that the downstream propagation and upstream propagation due to inefficiency in one agent's trade are different in nature. The downstream propagation (i.e. distortion in trades) is due to strategic complementarities. However, the upstream propagation is due to both complementarities and the need to reduce information rent.

Consider a market including a firm and three consumers connected via simple chain as depicted in Fig. 5. The payoff function of each agent is given as follows:

$$
\begin{aligned}
u_1(x_1, x_2, x_3, t_1) &= a x_1 - \frac{1}{2}x_1^2 - t_1, \\
u_2(\theta_2, x_1, x_2, x_3, t_2) &= a x_2 - \frac{1}{2}x_2^2 + \theta_2\ x_1\ x_2 - t_2, \\
u_3(x_1, x_2, x_3, t_3) &= a x_3 - \frac{1}{2}x_3^2 + \gamma\ x_2\ x_3 - t_3,
\end{aligned}
$$

where x_i and t_i are consumption and payment of agent i, respectively.

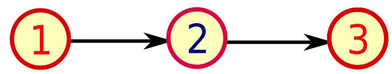

Figure 5: Line network. Interconnection among the agents.

In this market, agents 2 and 3 obtain positive externality from the consumptions of their neighbors. The externality from agent 2 to agent 3 is weighted by $\gamma \in (0\ 1)$ that is publicly known. In contrast, the externality from agent 1 to agent 2 is weighted by θ_2 that is agent 2's *private* information. It is, however, commonly known that $\theta_2 \in \{\underline{\theta}, \bar{\theta}\}$, $0 < \underline{\theta} < \bar{\theta} \leq \gamma < \frac{1}{\sqrt{2}}$, with $\text{Prob}\{\theta_2 = \underline{\theta}\} = \mathsf{v} > \frac{\Delta\theta}{\bar{\theta}}$, where $\Delta\theta \triangleq \bar{\theta} - \underline{\theta}$. Firm's objective is to devise a menu of optimal incentive quantity-price pairs $\{(x_i(\theta_2), t_i(\theta_2))\}$ for each agent i, given agent 2's report $\theta_2 \in \{\underline{\theta}, \bar{\theta}\}$, so as to maximize her total ex-ante profit subject to the corresponding participation constraint (PC) of each agent $i \in \{1, 2, 3\}$,[54] and incentive compatibility constraints (IC) of agent 2. Thus, firm's problem is precisely written as:

$$\begin{aligned} \max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^{3}} \quad & \mathsf{v} \sum_{i=1}^{3} t_i(\underline{\theta}) + (1 - \mathsf{v}) \sum_{i=1}^{3} t_i(\bar{\theta}) \\ \text{subject to} \quad & \text{PC}_i(\theta_2), \quad \theta_2 \in \{\underline{\theta}, \bar{\theta}\}, \text{ for all } i \in 1, 2, 3, \\ & \text{IC}_2(\theta_2), \quad \theta_2 \in \{\underline{\theta}, \bar{\theta}\}. \end{aligned} \tag{19}$$

In the full information case where θ_2 is commonly known, only PC constraints are active in the firm's problem, thus, the first-best trade profile maximizes the social surplus. In the second-best solution, the incentive constraint is active. Thus, in the firm's problem, the IC constraint of agent 2 binds at her efficient type and firm must give some rent to agent 2 for revealing her type truthfully. Importantly, given the above network structure, this rent is controlled not only by agent 2's trade but also agent 1's trade. To understand this better, let us first characterize the first and second-best solutions.

[54] The reservation utility of each agent is normalized to zero.

Lemma 2. *The first and second-best trade profiles are given by:*

$$x_1^{SB}(\bar{\theta}) = x_1^{FB}(\bar{\theta}), \quad x_1^{FB}(\underline{\theta}) = a + \underline{\theta}\, x_2^{FB}(\underline{\theta}), \quad x_1^{SB}(\underline{\theta}) = a + (\underline{\theta} - \mathsf{S})\, x_2^{SB}(\underline{\theta})$$

$$x_3^{SB}(\bar{\theta}) = x_3^{FB}(\bar{\theta}), \quad x_3^{FB}(\underline{\theta}) = a + \gamma\, x_2^{FB}(\underline{\theta}), \quad x_3^{SB}(\underline{\theta}) = a + \gamma\, x_2^{SB}(\underline{\theta})$$

$$x_2^{SB}(\bar{\theta}) = x_2^{FB}(\bar{\theta}), \quad x_2^{FB}(\underline{\theta}) = \frac{a(1 + (\underline{\theta} + \gamma))}{1 - (\underline{\theta}^2 + \gamma^2)}. \quad x_2^{SB}(\underline{\theta}) = \frac{a(1 + ((\underline{\theta} - \mathsf{S}) + \gamma))}{1 - ((\underline{\theta} - \mathsf{S})^2 + \gamma^2)}.$$

where $\mathsf{S} \triangleq \left(\frac{1-\mathsf{v}}{\mathsf{v}}\right)\Delta\theta > 0.$

Proof. See Appendix (omitted proofs). □

As expected, when agent 2 is efficient, i.e. $\theta_2 = \bar{\theta}$, the first and second-best are equal, i.e. $x_i^{SB}(\bar{\theta}) = x_i^{FB}(\bar{\theta})$, $i \in \{1, 2, 3\}$ and when agent 2 is inefficient, i.e. $\theta_2 = \underline{\theta}$, agent 2's allocation is distorted downward, i.e. $x_2^{SB}(\underline{\theta}) < x_2^{FB}(\underline{\theta})$. However, while agents 1's and 3's payoffs are both common knowledge, due to the positive network externality, in contrast with adverse selection environments with no externality,[55] distortion in agent 2's trade yields downward distortion in agents 1's and 3's trades. Most importantly, this distortion has *unequal* effects on agents 1's and 3's trades. To understand the reasoning, we look at agent 2's information rent, that is $R_2 = \Delta\theta\; x_1^{SB}(\underline{\theta})x_2^{SB}(\underline{\theta})$. Essentially, the downstream propagation and upstream propagations are different in nature. The downstream propagation (i.e. distortion in trades) is only because of *strategic complementarities*. Distortion in the allocation of the low-type agent 2 causes the optimal allocation of agent 3 to change. But this change is only there as a consequence of the distortion to low-type 2. And of course, this can be simply seen in the first order optimality condition for $x_3^{SB}(\underline{\theta})$: any change in $x_2^{SB}(\underline{\theta})$ has to be reflected in $x_3^{SB}(\underline{\theta})$. But, the same is *not* true for agent 1's allocation. Changing 1's trade also distort's high-type 2's *incentive* in mimicking low-type 2. In fact, by distorting 1's trade downward, the firm makes sure that 2's IC constraint is satisfied in *a cheaper way*, as essentially the effective type of agent 2 is now $\theta_2 x_1^{SB}(\underline{\theta})$. Lower $x_1^{SB}(\underline{\theta})$ helps the firm reduces the rent she has to pay to agent 2's high-type.

Example B-2 (Kite Network) Consider the following network:[56] Let us assume all

[55] Since the firm has complete information about agents 1's and 3's payoffs, as it is standard in adverse selection problems, one would expect to achieve *efficiency* in their trades, i.e., equality of the second and the first-bests.

[56] Parameters: $g_{ij} = g_{ji} \in \{0, .15\}$, $i, j \in \{1, 2, \cdots, 5\}$, capturing the interrelations among the agents. Let θ_i be uniformly distributed on $[\frac{2}{3}\ 1]$ (b is normalized to 1 and $a = 10$). In this network agent 2 is most central in terms of Bonacich centrality.

the agents report the efficient type $\bar{\theta}$ except agent 6. Agent 6's report varies from the lowest to the highest type, i.e., $\theta_i = \bar{\theta}$ for $1 \leq i \leq 5$ and $\underline{\theta} \leq \theta_6 \leq \bar{\theta}$. If there was no network there would be no distortion for the efficient agents $1, 2, 3, 4, 5$ and onlyagent 6's allocation would distort downward. However, due to strategic complementarities and the interconnection among the agents, distortion in agent 6's allocation propagates in the *whole* network. More importantly, each agent ,depending on its proximity to agent 6 experiences different amount of reduction in its allocation. The following figure shows the corresponding distortion and the way it unequally propagates throughout the network as θ_6 changes from $\underline{\theta}$ to $\bar{\theta}$. Note that when agent 6 is efficient, i.e. $\theta_6 = \bar{\theta}$, distortion vanishes, i.e. the first and second-best for all agents are equal. However, when $\theta_6 < \bar{\theta}$ distortion propagates and all agents experience reduction in their allocations, even though their types are *efficient*.

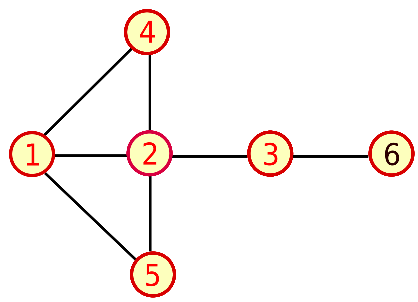

Figure 6: Kite network. Interconnection among the agents.

Example B-3 (different orders in bilateral and multilateral contracts) This example compares bilateral and multilateral contracts in a network with interconnections as in Example B-1 (above). We consider a market including a firm and three consumer agents. Fig. 5 visualizes the interconnection among the agents.

Let us first start with the multilateral contract.

Multilateral contract: The payoff function of each agent is as follows:

$$u_1(x_1, x_2, x_3, t_1) = ax_1 - \frac{1}{2}x_1^2 - t_1,$$
$$u_2(\theta_2, x_1, x_2, x_3, t_2) = ax_2 - \frac{1}{2}x_2^2 + \theta_2\; x_1\; x_2 - t_2,$$
$$u_3(x_1, x_2, x_3, t_3) = ax_3 - \frac{1}{2}x_3^2 + \bar{\theta}\; x_2\; x_3 - t_3.$$

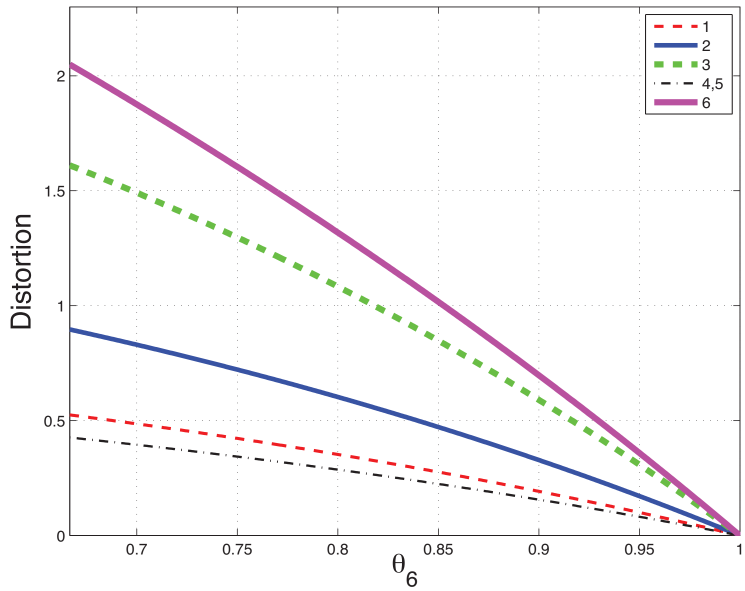

Figure 7: With decreasing θ_6 from $\bar{\theta} = 1$ (i.e., her efficient type) to $\underline{\theta} = \frac{2}{3}$, more distortion propagates throughout the network, reducing more other agents' allocations, given that agents $1, 2, \cdots, 5$ are all efficient. More importantly, agents $1, 2, \cdots, 5$ according to their locations in the network experience different reductions in their allocations.

In the above network, agents 2 and 3 obtain positive externality from the consumptions of their neighbors. The externality from agent 2 to agent 3 is weighed by $\bar{\theta}$ that is publicly known. However, the externality from agent 1 to agent 2 is weighed by θ_2 that is agent 2's *private* information. It is, however, commonly known that $\theta_2 \in \{\underline{\theta}, \bar{\theta}\}$, $0 < \underline{\theta} < \bar{\theta}$, with $\text{Prob}\{\theta_2 = \underline{\theta}\} = \nu > \frac{\Delta\theta}{\bar{\theta}}$, where $\Delta\theta \triangleq \bar{\theta} - \underline{\theta}$. Firm's objective is to devise a menu of optimal incentive quantity-price pairs $\{(x_i(\theta_2), t_i(\theta_2))\}$ for each agent i, given agent 2's report $\theta_2 \in \{\underline{\theta}, \bar{\theta}\}$, so as to maximize her total ex-ante profit subject to the corresponding PC of each agent $i \in \{1, 2, 3\}$, and IC constraint of agent 2. Therefore, the firm's problem is written as:

$$\begin{aligned} \max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^3} \quad & \nu \sum_{i=1}^{3} t_i(\underline{\theta}) + (1-\nu) \sum_{i=1}^{3} t_i(\bar{\theta}) \\ \text{subject to} \quad & \text{PC}_i(\theta_2), \quad \theta_2 \in \{\underline{\theta}, \bar{\theta}\}, \text{ for all } i \in 1, 2, 3, \\ & \text{IC}_2(\theta_2), \quad \theta_2 \in \{\underline{\theta}, \bar{\theta}\}. \end{aligned}$$

Solving the above program simply implies:

$$x_1^{SB}(\bar{\theta}) = x_3^{SB}(\bar{\theta}),$$
$$x_1^{SB}(\underline{\theta}) < x_3^{SB}(\underline{\theta}).$$

The above implies that (due to agent 2's information rent $R_2(\underline{\theta}) = \Delta\theta x_1(\underline{\theta})x_2(\underline{\theta})$) if agent 2 reports her efficient type, then agents 1 and 3's allocations are *exactly the same*. But agent 1's allocation is distorted (downward) more than 3's when agent 2 reports her inefficient type.

Next, we consider the bilateral contract in the same network

Bilateral contract: We show using bilateral contracts taht the above intuition does not carry over. Here, *independent* of agent 2's report, agent 1's allocation is *always* lower than agent 3's allocation. Focusing on the bilateral contract, each agent i's payoff is written as follows:

$$u_1(x_1, t_1) = ax_1 - \frac{1}{2}x_1^2 - t_1,$$
$$u_2(\theta_2, x_2, t_2) = ax_2 - \frac{1}{2}x_2^2 + \theta_2\ x_1\ x_2 - t_2, \quad \theta_2 \in \{\underline{\theta}, \bar{\theta}\}$$
$$u_3(x_3, t_3) = ax_3 - \frac{1}{2}x_3^2 + \bar{\theta}\ x_3\ \mathbf{E}[x_2(\theta_2)] - t_3,$$

and, thus, the firm's objective is to characterize

$$\{(x_1, t_1), \{(x_2(\underline{\theta}), t_2(\underline{\theta}))), (x_2(\bar{\theta}), t_2(\bar{\theta}))\}, (x_3, t_3)\},$$

maximizing her ex-ante profit with respect to the IC and PC constraints, Eqs. (11) and (12), respectively. Finally, we have the following.

Lemma 3. *Using bilateral contracts,* $x_3^{SB} - x_1^{SB} = \Delta\theta\ x_2^{SB}(\underline{\theta}) > 0.$

Therefore, in bilateral contracts independent of agent 2's report, $x_3^{SB} > x_1^{SB}$. This is because agent 2's information rent is $R_2(\underline{\theta}) = \Delta\theta\ x_1\ x_2(\underline{\theta})$, distorting (downward) more agent 1's allocation.

In summary, bilateral and multilateral contracts may induce different orders on the agents' allocations under presence of network externalities. By similar argument, one can show this effect can be even dramatic (for details, see Fig. 8). In the next section by characterizing the optimal bilateral trade profiles we present a more precise comparison for general networks.

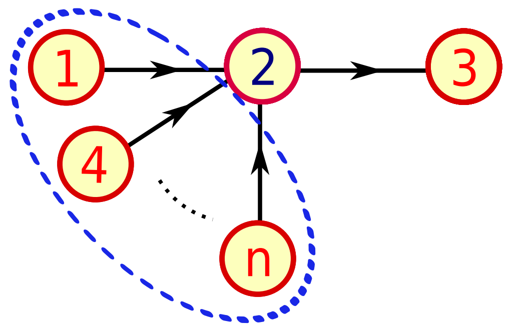

Figure 8: Star network. Using multilateral contracts, if agent 2 reports her efficient type, then $x_3(\bar{\theta}) = x_i(\bar{\theta})$, for all $i \in \{1, 4, 5, \cdots, n\}$. However, using bilateral contracts, if agent 2 reports her efficient type, then $x_3(\bar{\theta}) > x_i(\bar{\theta})$, for all $i \in \{1, 4, 5, \cdots, n\}$.

[**Analysis of Example B-1**] **Full information (First-best)**: Let θ_2 be commonly known. Hence, in the firm's problem only PC constraints are active, and, thus, the first-best trade profile is characterized as follows.

Lemma 4 (first-best)**.** *The first-best trade profile is given by:*

$$x_1^{FB}(\bar{\theta}) = a + \bar{\theta}\; x_2^{FB}(\bar{\theta}), \qquad x_1^{FB}(\underline{\theta}) = a + \underline{\theta}\; x_2^{FB}(\underline{\theta}),$$
$$x_3^{FB}(\bar{\theta}) = a + \gamma\; x_2^{FB}(\bar{\theta}), \qquad x_3^{FB}(\underline{\theta}) = a + \gamma\; x_2^{FB}(\underline{\theta}),$$
$$x_2^{FB}(\bar{\theta}) = \frac{a(1 + (\bar{\theta} + \gamma))}{1 - (\bar{\theta}^2 + \gamma^2)}, \qquad x_2^{FB}(\underline{\theta}) = \frac{a(1 + (\underline{\theta} + \gamma))}{1 - (\underline{\theta}^2 + \gamma^2)}.$$

We next characterize the second-best solution.

Sketch of the proof of Lemma 4. In the full information case, since the payoff functions of all agents are commonly known, the $PC_1(\underline{\theta}), PC_2(\underline{\theta}), PC_3(\underline{\theta}), PC_1(\bar{\theta}), PC_2(\bar{\theta})$, and $PC_3(\bar{\theta})$ must bind, characterizing $t_1(\underline{\theta})$, $t_2(\underline{\theta}), t_3(\underline{\theta})$, $t_1(\bar{\theta}), t_2(\bar{\theta})$, and $t_3(\bar{\theta})$, respectively. Then, plugging the characterized payments into the objective function and taking the first optimality condition yields the result.

□

Sketch of the proof of Lemma 2. Since the payoff functions of agents 1 and 3 are both fully known, thus the corresponding $PC_1(\underline{\theta}), PC_3(\underline{\theta}), PC_1(\bar{\theta})$ and $PC_3(\bar{\theta})$ must bind, characterizing $t_1(\underline{\theta})$, $t_3(\underline{\theta})$, $t_1(\bar{\theta})$, and $t_3(\bar{\theta})$, respectively. In addition $PC_2(\underline{\theta})$ must bind, characterizing $t_2(\underline{\theta})$. In addition, the (downward incentive constraint) $IC_2(\bar{\theta})$ must bind, characterizing $t_2(\bar{\theta})$, plus the corresponding information rent that is equal to $R = \Delta\theta x_1(\underline{\theta})x_2(\underline{\theta})$. The result then follows from the first order optimality condition of the objective function after plugging the payments, characterized in the above, into it. □

Proof of Lemma 3. [**Analysis of Example B-3**] For ease of exposition, let $\ell \triangleq x_2^{SB}(\underline{\theta})$ and $\hbar \triangleq x_2^{SB}(\bar{\theta})$, and ignore the superscript SB for the other agents. All the PC constraints for agents 1, 3, and 2 (at her inefficient type) as well as the IC of agent 2 at her efficient type must bind. Thus, the payment are as follows:

$$\begin{aligned}
t_1 &= ax_1 - \frac{1}{2}x_1^2, \\
t_3 &= ax_3 - \frac{1}{2}x_3^2 + \gamma\, x_3\, \mathbf{E}[x_2(\theta_2)], \\
t_2(\underline{\theta}) &= a\ell - \frac{1}{2}\ell^2 + \underline{\theta}\, \ell\, x_1, \\
t_2(\bar{\theta}) &= a\hbar - \frac{1}{2}\hbar^2 + \bar{\theta}\, \hbar\, x_1 - \underbrace{\Delta\theta\, x_1\, \ell}_{\text{information rent}},
\end{aligned}$$

Firm's objective becomes the following program:

$$\max_{x_1, x_2, \ell, \hbar \in \mathbb{R}_+} \quad t_1 + t_3 + \mathsf{v} t_2(\underline{\theta}) + (1-\mathsf{v})t_2(\bar{\theta})$$

By first order optimality condition, we have:

$$\begin{aligned}
x_1 &= a + \mathsf{v}\underline{\theta}\ell + (1-\mathsf{v})(\gamma\hbar - \Delta\theta\, \ell), \\
x_3 &= a + \gamma\, \mathbf{E}[x_2(\theta_2)], \\
\hbar &= a + \gamma(x_1 + x_3), \\
\ell &= a + \left(\underline{\theta} - \frac{1-\mathsf{v}}{\mathsf{v}}\Delta\theta\right)x_1 + \gamma x_3 > 0.
\end{aligned}$$

Therefore, $x_3 - x_1 = \ell\Delta\theta > 0$. □

C Role of influential agents in core-periphery structures: Infinite vs. Vanishing profit gap

In this section we formally prove the statements of Example 3 in Section 5.2.

In the next proposition, using Lemma 2, we show that among all networks within the class of star-family the maximum loss in the firm's profit occurs in the star-inward network.

Proposition 12. *Among all* $\mathsf{G} \in \mathfrak{G}(n)$ *(see Definition (5)),* star-inward *maximizes* $\Pi_{\mathsf{G}}^{multi.} - \Pi_{\mathsf{G}}^{bi.}$.

Proof. See the Appendix—proof section. □

As the figure in Example 3 shows, the maximum loss (within the star-family) occurs in star-inwards. Moreover, the figure highlights the fact that with increasing number of agents, the firm's loss for using bilateral contracts grows unboundedly. This, of course, is not always true and depends on the network constructions. The following propositions highlight the points.

Proposition 13. *Consider the sequence of star-inward networks* $\{\mathsf{G}(n)\}_{n=2}^{\infty}$, *where* $\alpha_n \triangleq g_{12} = g_{13} = \cdots = g_{1n}$, *and the rest of entries are zero. If* $\alpha_n = \frac{1}{\sqrt{(n-1)(\mu^2+\sigma^2+\epsilon)}}$, *where* $\epsilon > 0$ *is constant, then as n grows:*

$$\Pi_{\mathsf{G}(n)}^{multi.} - \Pi_{\mathsf{G}(n)}^{bi.} \longrightarrow \infty.$$

Proof. See the Appendix—proof section. □

The above proposition shows that the marginal benefit of using multilateral contracts may grow unboundedly in networks that have major asymmetry in their in-degrees. When there is an agent dominant in its in-degree (e.g., star-inward) then the loss in the firm's profit because of using the simpler bilateral contracts may become unbounded. In summery, in large networks exhibiting large asymmetry in their in-degrees, firm's restriction to the simpler bilateral contracts may result in major loss in firm's profit.

When will the profit gap vanish? In contrast, the following proposition asserts that in large directed core-periphery economies the difference (in profit) between multilateral and bilateral contracts might be negligible. This is intuitive because, as shown in Proposition 8 (part *(ii)*), decreasing the influencing weights (i.e., $g_{ij}\ \forall i,j$) reduces the benefit of using multilateral contracts. To prove the result we use Proposition 12, that is, in the class of star-like core-periphery structures[57] the benefit of using multilateral contacts in star-inward networks is higher than the others.

Proposition 14. *Let $\mathcal{G}(n)$ denote the set of all economies including n agents with the directed star network[58] structure. For any $\mathsf{G}(n) \in \mathcal{G}(n)$, if its non-zero weights are equal to $\beta_n \triangleq \frac{1}{(n-1)\sqrt{\mu^2+\sigma^2+\epsilon}}$, where $\epsilon > 0$ is constant, then as n grows:*

$$\Pi^{multi.}_{\mathsf{G}(n)} - \Pi^{bi.}_{\mathsf{G}(n)} \longrightarrow 0.$$

Proof. See the Appendix—proof section. □

D Benefits of uncertainty

Given the findings in Proposition 3 and Proposition 5, it is important to distinguish between two sources of uncertainty in the model. One source is the uncertainty among the agents, and the other is the uncertainty between the firm and each agent. Under the imperfect information case (the second-best) both sources are present. Under the full information case (the first-best) only uncertainty among the agents is present. Importantly, the uncertainty between the firm and each agent manifests its effect in the disparity of the first-best and the second-best solutions in which, as in a canonical adverse selection problem, θ changes to $\psi(\theta)$, which is *not* structural. However, the impact of uncertainty among the agents is more profound and is captured by the *structure* of the first-best solution. Similar arguments hold in the multilateral contract. The following remark summarizes.

Remark 4. *We note that similar to multilateral contracts, analyzing the effect of uncertainty in the full information case and imperfect information case are essentially the same. In the imperfect information case the uncertainty is in the virtual type, whereas in the full information case uncertainty is in the type θ.*

[57] In which $g_{ij} \in \{0, \kappa\}$ for all i, j, $(\kappa > 0)$.
[58] With no parallel links.

We next consider the impact of uncertainty on the firm's expected profit.

Proposition 15. *The optimal second-best expected profit of the firm increases as the uncertainty in the agents' (virtual) types increases.*

Proof. See the Appendix—proof section. □

The preceding proposition states that the *firm prefers the uncertainty in agents' types to be higher*. It is helpful to examine the reasoning behind this statement. As mentioned before, increasing the uncertainty in the agents' types has two effects: it not only increases the adverse selection effect that the firm faces against each agent (uncertainty between firm and agents), but also increases the uncertainty in the beliefs of one agent regarding her neighbors' types (i.e., uncertainty among the agents). The former effect tends to reduce the firms expected profit, as evidenced by the fact that the firm would obtain higher profits in the full information setting. However, from Proposition 4, we see that higher uncertainty about a neighbor's type increases the quantity consumed by other agents, and hence the latter effect tends to increase the firm's expected profit. Thus, there are two opposing forces at play here and the proposition states that the latter effect is stronger. We attribute this to the positive externality in the network: increasing the consumption profile of any one agent causes consumption of other agents in the network to increase thus earning even higher profits for the firm.[59]

The following corollary shows that in the first-best allocation, firm's profit is increasing in the variance of the types.

Corollary 2. *Let μ be fixed. Then, firm's first-best expected profit increases with a greater uncertainty among the agents, i.e.* $\frac{\partial \Pi_G^{bi.}}{\partial \sigma^2} > 0$.[60]

Proof. See the Appendix—proof section. □

[59]It is also worth noting that employing Taylor expansions for the moments of the function $\psi(\cdot)$ of the random variable θ_i we have: $Var(\psi(\tilde{\theta}_i)) \propto \left(\frac{d}{d\rho}\psi(\rho)|_{\rho=\mathbf{E}[\tilde{\theta}_i]}\right)^2 \ Var(\tilde{\theta}_i) = (\dot{\psi}(\mathbf{E}[\tilde{\theta}_i]))^2 \ Var(\tilde{\theta}_i)$. That is, increasing $Var(\tilde{\theta}_i)$ increases $Var(\psi(\tilde{\theta}_i))$, when θ is around its mean.

[60]One way to change σ^2 and keep μ fixed is via changing $\Delta\theta = \bar{\theta} - \underline{\theta}$. That is, to ensure that a change in σ^2 does not affect μ, we change $\Delta\theta$ so that μ is kept fixed.

E Appendix: Omitted Proofs

Proof of Proposition 1. We prove each part of the proposition separately:

Proof of Part (i): To find $\{x_i^{FB}(\cdot)\}_{i=1}^n$, we maximize the objective function in (5) point-wise for any type profile θ. Thus, for a given θ, to find $(x_1^{FB}(\theta), x_2^{FB}(\theta), \cdots, x_n^{FB}(\theta))$ we faced with the following program:

$$\max_x \sum_{i=1}^n \left[a x_i - \frac{b}{2} x_i^2 + x_i \theta_i \sum_{j=1}^n g_{ij} x_j \right]$$
$$\text{subject to:} \quad x_i \geq 0 \quad \forall i \in V. \tag{20}$$

The Hessian matrix corresponding to the objective function in (20) is given by:

$$\begin{pmatrix} -b & \theta_1 g_{12} + \theta_2 g_{21} & \cdots & \theta_1 g_{1n} + \theta_n g_{n1} \\ \theta_1 g_{12} + \theta_2 g_{21} & -b & \cdots & \theta_2 g_{2n} + \theta_n g_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1 g_{1n} + \theta_n g_{n1} & \theta_2 g_{2n} + \theta_n g_{n2} & \cdots & -b \end{pmatrix}.$$

The Hessian matrix is a Hermitian, strictly diagonally dominant matrix[61] (due to Assumption 1) with real negative diagonal entries, thus it is negative semi-definte (due to Sylvester's criterion, Horn and Johnson [2012]). As a result, the objective function in (20) is concave.

Next, assuming the solution to program (20) is interior implies that it satisfies the following first order optimality condition:

$$a - b x_i^{FB}(\theta) + \left(\theta_i \sum_{j=1}^n g_{ij} x_j^{FB}(\theta) + \sum_{j=1}^n g_{ji} \theta_j x_j^{FB}(\theta) \right) = 0, \qquad \forall\ i \in V. \tag{21}$$

Recall that $\mathsf{M}_\theta = \text{diag}(\theta_1, \theta_2, \cdots, \theta_n)$. Thus, the above equation in its matrix form is written

[61] A matrix is called strictly diagonally dominant if for every row of the matrix, the magnitude of the diagonal entry in a row is larger than the sum of the magnitudes of all the other (non-diagonal) entries in that row. That is, the matrix $A = [a_{ij}]_{(i,j) \in V^2}$ is diagonally dominant if $|a_{ii}| > \sum_{j \neq i} (|a_{ij}| + |a_{ji}|)$, for all $i \in V$.

as:

$$a\mathbf{1} + (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta)\mathbf{x}^{FB}(\theta) = b\mathbf{x}^{FB}(\theta),$$

Hence, we obtain:

$$\mathbf{x}^{FB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta)]^{-1}\mathbf{1}.$$

Note that since $(b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta))$ is a strictly diagonally dominant matrix, it is invertible (due to Levy-Desplanques theorem, Horn and Johnson [2012]). Furthermore, notice that $x^{FB}(\theta)$ is in the positive orthant, since $a > 0$ and Assumption 1. Finally, we show that there is no corner solution, and, consequently, the above solution is unique. Suppose it were not true; that is, there exists a non-interior solution that we denote it by y^*. Let $W \subset V$, $(W \neq \emptyset)$, such that $x_i^* = 0$ if and only if $i \in V/W$ (that is, $x_i^* > 0$ when $i \in W$). Since $a > 0$, thus $\mathbf{0}$ can not be an optimal solution, therefore, $W \neq \emptyset$. Let $y_i^* = 0$, due to the optimality condition, by (21) we should have:

$$a - by_i + \left(\theta_i \sum_{j=1}^n g_{ij} y_j + \sum_{j=1}^n g_{ji}\theta_j y_j\right)\Bigg|_{y=y^*} = a + \left(\theta_i \sum_{j=1}^n g_{ij} y_j^* + \sum_{j=1}^n g_{ji}\theta_j y_j^*\right) \leq 0.$$

However, since $a > 0$, achieving the last inequality is impossible, which is a contradiction.

Proof of Part (ii): Define, for every i and $\theta_i \in [\underline{\theta}\ \bar{\theta}]$,

$$V_i(\theta_i) \triangleq \mathbf{E}_{\theta_{-i}}\left[a x_i(\theta_i, \theta_{-i}) - \frac{b}{2} x_i^2(\theta_i, \theta_{-i})\right] \tag{22}$$

$$\gamma_i(\theta_i) \triangleq \mathbf{E}_{\theta_{-i}}\left[x_i(\theta_i, \theta_{-i}) \sum_{j \neq i} g_{ij} x_j(\theta_i, \theta_{-i})\right] \tag{23}$$

$$T_i(\theta_i) \triangleq \mathbf{E}_{\theta_{-i}}\left[t_i(\theta_i, \theta_{-i})\right]. \tag{24}$$

Thus, the agent i's interim utility for reporting $\hat{\theta}_i$, while her real type is θ_i, is given by:

$$U_i(\theta_i, \hat{\theta}_i) = \theta_i \gamma_i(\hat{\theta}_i) + V_i(\hat{\theta}_i) - T_i(\hat{\theta}_i).$$

By appealing to the revenue equivalence theorem, a direct quantity-price schedule

$\{x_i^{SB}(\cdot), t_i^{SB}(\cdot)\}_{i=1}^n$ is IC and PC if and only if for every $i \in V$:

(i) $\gamma_i(\theta_i)$ is increasing;

(ii) For every $\theta_i \in [\underline{\theta}\ \bar{\theta}]$:

$$T_i(\theta_i) = \theta_i \gamma_i(\theta_i) + V_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau. \tag{25}$$

Given the specification of the incentive compatible and individual rational mechanisms the firm's problem is re-written as:

$$\max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^n} \quad \sum_{i=1}^n \mathbf{E}_{\theta_i}\Big[T_i(\theta_i)\Big]$$

$$\text{subject to} \quad T_i(\theta_i) = \theta_i \gamma_i(\theta_i) + V_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau \quad \forall i,\ \theta_i, \tag{26}$$

$$\gamma_i(\theta_i) \text{ is increasing} \quad \forall\ i \tag{27}$$

To find the solution of the above program we first ignore the monotonicity constraints (27), later we show (27) is indeed satisfied. Thus, plugging (26) into the objective function, and using the fact that $\mathbf{E}_{\theta_i}\left[\int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau\right] = \mathbf{E}_{\theta_i}\left[\frac{1-F(\theta_i)}{f(\theta_i)}\gamma_i(\theta_i)\right]$ we obtain

$$\max_{\{x_i(\cdot)\}_{i=1}^n} \quad \sum_{i=1}^n \mathbf{E}_{\theta_i}\left[\psi(\theta_i)\gamma_i(\theta_i) + V_i(\theta_i)\right], \tag{28}$$

recall that $\psi(\theta_i) = \theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}$ (virtual type). Given the definitions in (22)-(24), and the fact that $\mathbf{E}[\cdot]$ is a linear operator, (28) can be re-written as:

$$\max_{\{x_i(\cdot)\}_{i=1}^n} \quad \mathbf{E}_\theta \sum_{i=1}^n \left[\psi(\theta_i) x_i(\theta_i, \theta_{-i}) \sum_{j \neq i} g_{ij} x_j(\theta_i, \theta_{-i}) + a x_i(\theta_i, \theta_{-i}) - \frac{b}{2} x_i^2(\theta_i, \theta_{-i})\right]$$

$$\equiv \max_{\{x_i(\cdot)\}_{i=1}^n} \quad \mathbf{E}_\theta \sum_{i=1}^n \left[\psi(\theta_i) x_i(\theta) \sum_{j \neq i} g_{ij} x_j(\theta) + a x_i(\theta) - \frac{b}{2} x_i^2(\theta)\right]. \tag{29}$$

To find the optimal solution to (29), we maximize it point-wise. Let $\theta \in [\underline{\theta}\ \bar{\theta}]^n$ be fixed and given, hence $\{x_i^{SB}(\theta)\}_{i=1}^n$ solves the following program:

$$\max_{\{x_i(\theta)\}_{i=1}^n} \sum_{i=1}^n \left[\psi(\theta_i) x_i(\theta) \sum_{j \neq i} g_{ij} x_j(\theta) + a x_i(\theta) - \frac{b}{2} x_i^2(\theta) \right]. \tag{30}$$

Note that the objective function in (29) is concave. Because its corresponding Hessian matrix,

$$\begin{pmatrix} -b & \psi(\theta_1)g_{12} + \psi(\theta_2)g_{21} & \cdots & \psi(\theta_1)g_{1n} + \psi(\theta_n)g_{n1} \\ \psi(\theta_1)g_{12} + \psi(\theta_2)g_{21} & -b & \cdots & \psi(\theta_2)g_{2n} + \psi(\theta_n)g_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \psi(\theta_1)g_{1n} + \psi(\theta_n)g_{n1} & \psi(\theta_2)g_{2n} + \psi(\theta_n)g_{n2} & \cdots & -b \end{pmatrix},$$

is Hermitian and strictly diagonally dominant and thus it is negative semi-definte (due to Sylvester's criterion, Horn and Johnson [2012]). The diagonally dominant property of the above matrix is ensured by Assumption 1, 2, and the fact that $\psi(\theta_i) \leq \theta_i$ for all $\theta_i \in [\underline{\theta}\ \bar{\theta}]$ and $i \in V$.

The first order optimality condition of (30) (assuming the solution is interior) yields

$$a - b x_i^{SB}(\theta) + \left(\psi(\theta_i) \sum_{j=1}^n g_{ij} x_j^{SB}(\theta) + \sum_{j=1}^n g_{ji} \psi(\theta_j) x_j^{SB}(\theta) \right) = 0, \qquad \forall\ i \in V. \tag{31}$$

Recall that $\mathsf{M}_\psi = \operatorname{diag}(\psi(\theta_1), \psi(\theta_2), \cdots, \psi(\theta_n))$. Thus, the above equation in its matrix form is written as:

$$a\mathbf{1} + (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)\mathbf{x}^{SB}(\theta) = b\mathbf{x}^{SB}(\theta).$$

Since $(b\mathsf{I} - (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi))$ is a strictly diagonally dominant matrix, it is invertible (due to Levy-Desplanques theorem, Horn and Johnson [2012]). Thus,

$$\mathbf{x}^{SB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)]^{-1}\mathbf{1}. \tag{32}$$

Furthermore, notice that $x^{SB}(\theta)$ is in the positive orthant, since $a > 0$ and Assumption 1 and 2.

Notice that, since $\psi(\theta_i) \geq 0$ for all $\theta_i \in [\underline{\theta}\ \bar{\theta}]$ and $i \in V$, similar to the proof of the first-best case, one can easily show $\mathbf{x}^{SB}(\theta)$, characterized in (32), is indeed the unique to the program (30), for any given θ.

Next, we present a lemma that we make a use of it in the proof later.

Lemma 5. *Let* $\mathsf{K} \triangleq [b\mathsf{I} - (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)]^{-1}$. *Then,* $\frac{\partial \mathsf{K}}{\partial \theta_i}$ *is a matrix with non-negative entries, for any* $i \in V$.

Proof. Observe that, by chain rule,

$$\mathbf{0} = \frac{\partial I}{\partial \theta_i} = \frac{\partial \mathsf{K}\mathsf{K}^{-1}}{\partial \theta_i} = \frac{\partial \mathsf{K}}{\partial \theta_i}\mathsf{K}^{-1} + \mathsf{K}\frac{\partial \mathsf{K}^{-1}}{\partial \theta_i}. \tag{33}$$

Furthermore, as G does not depend on θ_i, we have

$$\begin{aligned}\frac{\partial \mathsf{K}^{-1}}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i}\Big[b\mathsf{I} - (\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)\Big] = -\frac{\partial}{\partial \theta_i}\Big[(\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi)\Big] = -\frac{\partial \mathsf{M}_\psi}{\partial \theta_i}\mathsf{G} - \mathsf{G}^T\frac{\partial \mathsf{M}_\psi}{\partial \theta_i}\\ &= -(\mathsf{E}_{ii}\mathsf{G} + \mathsf{G}^T \mathsf{E}_{ii}),\end{aligned}$$

where $\mathsf{E}_{ii} \triangleq \frac{\partial \mathsf{M}_\psi}{\partial \theta_i}$ is a matrix with $\frac{\partial \psi(\theta_i)}{\partial \theta_i}$ at the $ii^{\,th}$ entry, and zero otherwise. Notice that $\frac{\partial \psi(\theta_i)}{\partial \theta_i} \geq 0$, (due to the monotonicty assumption in the hazard rate). Thus, using (57) we obtain

$$\frac{\partial \mathsf{K}}{\partial \theta_i} = -\mathsf{K}\frac{\partial \mathsf{K}^{-1}}{\partial \theta_i}\mathsf{K} = \mathsf{K}(\mathsf{E}_{ii}\mathsf{G} + \mathsf{G}^T \mathsf{E}_{ii})\mathsf{K}.$$

Thus, since the right hand side in the above equation is non-negative, $\frac{\partial \mathsf{K}}{\partial \theta_i}$ is a matrix with non-negative entries. □

This Lemma has two important implications: (i) $\frac{\partial x_i^{SB}(\theta)}{\partial \theta_i} \geq 0$, this property is intuitively immediate due to the IC constraint. (ii) $\frac{\partial x_j^{SB}(\theta)}{\partial \theta_i} \geq 0$, $(j \neq i)$, this property is due to the strategic complement property.

Finally, to wrap up the proof, it is left to show that $\gamma_i(\theta_i)$ is increasing in θ_i, for all $\theta_i \in [\underline{\theta}\ \bar{\theta}]$, and $i \in V$ (recall that monotonicity of $\gamma_i(\theta_i)$ is the constraint must be satisfied to achieve IC).

By the definition of $\gamma_i(\theta_i)$ (see (23)) and the above two points ((i) and (ii)), we have:

$$\begin{aligned}
\frac{\partial \gamma_i(\theta_i)}{\partial \theta_i} &= \mathbf{E}_{\theta_{-i}}\left[\frac{\partial\{x_i(\theta_i,\theta_{-i})\sum_{j\neq i} g_{ij}x_j(\theta_i,\theta_{-i})\}}{\partial \theta_i}\right] \\
&= \mathbf{E}_{\theta_{-i}}\left[\underbrace{\frac{\partial x_i(\theta_i,\theta_{-i})}{\partial \theta_i}}_{\geq 0 \text{ by (i)}}\sum_{j\neq i} g_{ij}x_j(\theta_i,\theta_{-i}) + x_i(\theta_i,\theta_{-i})\sum_{j\neq i} g_{ij}\underbrace{\frac{\partial x_j(\theta_i,\theta_{-i})}{\partial \theta_i}}_{\geq 0 \text{ by (ii)}}\right] \\
&\geq 0.
\end{aligned}$$

The proof is complete. □

Proof of Lemma 1. Proof of Part (i) follows directly from Proposition 1 and the definition of Bonacich centrality measure (see Def. 1). We proceed to prove the rest.

Given the definition of distortion vector and make a use of Proposition 1, we obtain:

$$\begin{aligned}
\mathbf{d}(\theta) &= a\Big([b\mathsf{I} - (\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1} - [b\mathsf{I} - (\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)]^{-1}\Big)\mathbf{1} \\
&= \frac{a}{b}\Big([\mathsf{I} - \frac{1}{b}(\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1} - [\mathsf{I} - \frac{1}{b}(\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)]^{-1}\Big)\mathbf{1}.
\end{aligned}$$

Next, define $M \triangleq [\mathsf{I} - \frac{1}{b}(\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1}$ and $N \triangleq [\mathsf{I} - \frac{1}{b}(\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)]^{-1}$. Thus, their entries are characterized as:

$$\begin{aligned}
m_{ij} &= \sum_{k=0}^{\infty}\frac{1}{b^k}f_{ij}^{[k]} \\
n_{ij} &= \sum_{k=0}^{\infty}\frac{1}{b^k}h_{ij}^{[k]}
\end{aligned}$$

where $f_{ij}^{[k]}$ is the ij entry of $(\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)^k$, and $h_{ij}^{[k]}$ is the ij entry of $(\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)^k$. Recall that $\psi(\tau) = \tau - \frac{1-F(\tau)}{f(\tau)}$, $\tau \in [\underline{\theta}\ \bar{\theta}]$. Thus, $\psi(\tau) \leq \tau$ with equality only at $\tau = \bar{\tau}$. This implies that $f_{ij}^{[k]} \geq h_{ij}^{[k]}$, for any k. Consequently, $m_{ij} \geq n_{ij}$, for all i,j. Since $\theta_j < \bar{\theta}$, thus $\psi(\theta_j) < \theta_j$. Consequently, the result follows. □

Proof of Proposition 2. We prove the parts separately as follows.

Part (i): For ease of exposition we consider agent 1, i.e., $i = 1$. Thus, $\theta_1 \leq \bar{\theta}$, while all the other agents are efficient, i.e., $\theta_j = \bar{\theta}$ (for all $j \neq 1$). Let $(\theta_1, \bar{\theta}_{-1})$ denote the corresponding type profile.

Since θ_1 is sufficiently close to $\bar{\theta}$, thus:

$$\mathcal{T}_1(\theta_1, \bar{\theta}_{-1}) = \mathcal{T}_1(\bar{\theta}_1, \bar{\theta}_{-1}) - (\bar{\theta} - \theta_1)\mathcal{T}_1'(\bar{\theta}_1, \bar{\theta}_{-1}), \tag{34}$$

where $\mathcal{T}_1'(\bar{\theta}_1, \bar{\theta}_{-1}) = \frac{\partial \mathcal{T}_1'(\theta_1, \bar{\theta}_{-1})}{\partial \theta_1}|_{\theta_1 = \bar{\theta}_1}$.

Notice that, by Definition 3, $\mathcal{T}_1(\bar{\theta}_1, \bar{\theta}_{-1}) = \mathbf{1}^T \mathbf{d}(\bar{\theta}_1, \bar{\theta}_{-1})$. According to Proposition (1), $\mathbf{d}(\bar{\theta}_1, \bar{\theta}_{-1}) = 0$, since $\mathsf{M}_\theta = \mathsf{M}_\psi = \mathrm{diag}(\bar{\theta}, \bar{\theta}, \cdots, \bar{\theta})$, and thus $\mathcal{T}_1(\bar{\theta}_1, \bar{\theta}_{-1}) = 0$.

Therefore, we only need to characterize $(\bar{\theta} - \theta_1)\mathcal{T}_1'(\bar{\theta}_1, \bar{\theta}_{-1})$. To do so, for ease of exposition, define $S^{-1} \triangleq (\mathsf{I} - \frac{1}{b}(\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta))^{-1}$ and $T^{-1} \triangleq (\mathsf{I} - \frac{1}{b}(\mathsf{M}_\psi \mathsf{G} + \mathsf{G}^T \mathsf{M}_\psi))^{-1}$. Therefore, due to Proposition 1, $\mathbf{d}(\theta) = \frac{a}{b}(S^{-1} - T^{-1})\mathbf{1}$.

Next, since $SS^{-1} = \mathsf{I}$, we have, $\frac{\partial S}{\partial \theta_1} S^{-1} + S \frac{\partial S^{-1}}{\partial \theta_1} = 0$, that yields:

$$\frac{\partial S^{-1}}{\partial \theta_1} = -S^{-1} \frac{\partial S}{\partial \theta_1} S^{-1}.$$

By the definition,

$$\begin{aligned} \frac{\partial S}{\partial \theta_1}|_{\theta_1 = \bar{\theta}} &= -\frac{1}{b}[\mathrm{diag}(1, 0, \cdots, 0)\mathsf{G} + \mathsf{G}^T \mathrm{diag}(1, 0, \cdots, 0)] \\ &= -\frac{1}{b}\left(R_1 + R_1^T\right), \end{aligned}$$

where,

$$R_1 \triangleq \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{35}$$

Similarly,

$$\begin{aligned} \frac{\partial T}{\partial \theta_1}|_{\theta_1 = \bar{\theta}} &= -\frac{1}{b}[\mathrm{diag}(\psi'(\bar{\theta}), 0, \cdots, 0)\mathsf{G} + \mathsf{G}^T \mathrm{diag}(\psi'(\bar{\theta}), 0, \cdots, 0)] \\ &= -\frac{1}{b}\psi'(\bar{\theta})\left(R_1 + R_1^T\right), \end{aligned}$$

where $\psi'(\bar{\theta}) = 1 - \phi'(\bar{\theta})$. Moreover, since $\psi(\bar{\theta}) = \bar{\theta}$, hence $S^{-1}|_{\theta=\bar{\theta}} = T^{-1}|_{\theta=\bar{\theta}}$. Therefore,

$$\begin{aligned}
\frac{\partial \mathbf{d}(\theta_i, \bar{\theta}_{-1})}{\partial \theta_1}|_{\theta_1=\bar{\theta}} &= \frac{a}{b}\left[-S^{-1}\frac{\partial S}{\partial \theta_1}S^{-1} + T^{-1}\frac{\partial T}{\partial \theta_1}T^{-1}\right]\mathbf{1} = \frac{a}{b}\left[-S^{-1}\left(\frac{\partial S}{\partial \theta_1} - \frac{\partial T}{\partial \theta_1}\right)S^{-1}\right]\mathbf{1} \\
&= \frac{a}{b^2}\left[-S^{-1}\Big((-1+\psi'(\bar{\theta}))(R_1 + R_1^T))\Big)S^{-1}\right]\mathbf{1} \\
&= \frac{a}{b^2}(1-\psi'(\bar{\theta}))\left[S^{-1}\Big(R_1 + R_1^T\Big)S^{-1}\right]\mathbf{1} \\
&= \frac{a}{b^2}\phi'(\bar{\theta})\left[S^{-1}\Big(R_1 + R_1^T\Big)S^{-1}\right]\mathbf{1}
\end{aligned}$$

The above equality along with (34) yields

$$\begin{aligned}
\mathcal{T}_1(\theta_1, \bar{\theta}_{-1}) &= -(\bar{\theta} - \theta_1)\mathcal{T}_1'(\bar{\theta}_1, \bar{\theta}_{-1}) \\
&= (\theta_1 - \bar{\theta})\ \mathbf{1}^T\ \frac{\partial \mathbf{d}(\theta_i, \bar{\theta}_{-1})}{\partial \theta_1}|_{\theta_1=\bar{\theta}} \\
&= \frac{a}{b^2}\Big((\theta_1 - \bar{\theta})\phi'(\bar{\theta})\Big)\mathbf{1}^T\left[S^{-1}\Big(R_1 + R_1^T\Big)S^{-1}\right]|_{\theta_1=\bar{\theta}}\ \mathbf{1} \\
&\overset{(a)}{=} \frac{a}{b^2}\Big((\theta_1 - \bar{\theta})\phi'(\bar{\theta})\Big)[k_1\ k_2\ \cdots k_n]\Big(R_1 + R_1^T\Big)[k_1\ k_2\ \cdots k_n]^T \\
&\overset{(b)}{=} \frac{a}{b^2}\Big((\theta_1 - \bar{\theta})\phi'(\bar{\theta})\Big)\left[2\sum_j k_1 k_j g_{1j}\right],
\end{aligned}$$

where (a) follows since $S^{-1}|_{\theta=\bar{\theta}}\mathbf{1} = T^{-1}|_{\theta=\bar{\theta}}\mathbf{1} = \Big(\mathsf{I} - \frac{1}{b}\bar{\theta}(\mathsf{G} + \mathsf{G}^T)\Big)^{-1}\mathbf{1} \triangleq [k_1\ k_2\ \cdots\ k_n]^T$, where k_i is agent i's Bonacich centrality in $\mathsf{G} + \mathsf{G}^T$. And (b) follows since[62]

$$\begin{aligned}
[k_1\ k_2\ \cdots k_n]\Big(R_1 + R_1^T\Big)[k_1\ k_2\ \cdots k_n]^T &= [k_1\ k_2\ \cdots k_n]\left[g_{11}k_1 + \sum_j g_{1j}k_j\ \ g_{12}k_1\ \ g_{13}k_1\ \ \cdots\ \ g_{1n}k_1\right]^T \\
&= k_1\sum_j g_{1j}k_j + \sum_j g_{1j}k_j k_1 \\
&= 2\sum_j g_{1j}k_j k_1.
\end{aligned}$$

[62] Recall that $g_{11} = 0$.

The above computations work for all i. Thus

$$\mathcal{T}_i(\theta_i, \bar{\theta}_{-i}) = \frac{2a}{b^2}(\theta_i - \bar{\theta})\phi'(\bar{\theta})\left[\sum_j k_i k_j g_{ij}\right]$$
$$= \frac{2a}{b^2}\,(\bar{\theta} - \theta_i)\,|\phi'(\bar{\theta})|\,\left[\sum_j k_i k_j g_{ij}\right]. \tag{36}$$

Notice that $\phi'(\bar{\theta}) < 0$.

Part (ii): Lets first derive $\mathbf{E}[\Pi^{FB}]$. As shown in Proposition 1, the optimal (first-best) trade profile is $\mathbf{x}^{FB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta)]^{-1}\mathbf{1}$. Thus, for each type profile θ the (ex post) firm's profit is given by the following. For ease of exposition, let x_i denote agent i's (first-best) allocation with respect to the type profile θ then

$$\sum_{i=1}^{n}\left(ax_i - \frac{b}{2}x_i^2 + x_i\theta_i \sum_j g_{ij}x_j\right) = a\mathbf{1}^T\mathbf{x} - \frac{b}{2}\mathbf{x}^T\mathbf{x} + (\mathsf{M}_\theta \mathbf{x})^T \mathsf{G}\mathbf{x}$$
$$= a\mathbf{1}^T\mathbf{x} - \frac{b}{2}\mathbf{x}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T(\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)\mathbf{x}$$
$$= a\mathbf{1}^T\mathbf{x} - \frac{1}{2}\left[\mathbf{x}^T\ \ [b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]\ \ \mathbf{x}\right]$$
$$= \frac{a^2}{2}\mathbf{1}^T[b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1}\mathbf{1}.$$

Therefore $\mathbf{E}[\Pi^{FB}] = \frac{a^2}{2}\mathbf{1}^T\mathbf{E}\left[[b\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1}\right]\mathbf{1}$.

Next, to derive $\mathbf{E}[\Pi^{SB}]$ we first note that as shown in Proposition 1 (see (26), (28), (29)) $\mathbf{E}[\Pi^{SB}] = \mathbf{E}\sum_{i=1}^n T(\theta_i) = \mathbf{E}\sum_{i=1}^n \left[\psi(\theta_i)y_i(\theta)\sum_{j\neq i} g_{ij}y_j(\theta) + ay_i(\theta) - \frac{b}{2}y_i^2(\theta)\right]$, where $y_i(\theta) = x_i^{SB}(\theta)$. Thus, similar to the full information case, let y_i denote agent i's (second-best) allocation with respect to the type profile θ then

$$\sum_{i=1}^{n}\left(ay_i - \frac{b}{2}y_i^2 + y_i\psi(\theta_i)\sum_j g_{ij}y_j\right) = a\mathbf{1}^T\mathbf{y} - \frac{b}{2}\mathbf{y}^T\mathbf{y} + (\mathsf{M}_\psi\mathbf{y})^T\mathsf{G}\mathbf{y}$$
$$= a\mathbf{1}^T\mathbf{y} - \frac{b}{2}\mathbf{y}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T(\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)\mathbf{y}$$
$$= a\mathbf{1}^T\mathbf{y} - \frac{1}{2}\left[\mathbf{y}^T \quad [b\mathsf{I} - (\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)] \quad \mathbf{y}\right]$$
$$= \frac{a^2}{2}\mathbf{1}^T[b\mathsf{I} - (\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)]^{-1}\mathbf{1}.$$

Thus,

$$\mathbf{E}[\Pi^{FB} - \Pi^{SB}] = \frac{a^2}{2}\mathbf{E}\mathbf{1}^T\left[[b\mathsf{I} - (\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1} - [b\mathsf{I} - (\mathsf{M}_\psi\mathsf{G} + \mathsf{G}^T\mathsf{M}_\psi)]^{-1}\right]\mathbf{1}$$
$$= \frac{a^2}{2}\mathbf{E}\,\mathcal{T}(\theta_i, \theta_{-i}).$$

Next, under the assumption that it is known to the firm that $\theta_j = \bar{\theta}$, for all $j \neq i$. Then, when $\theta_i \in \{\underline{\theta}, \bar{\theta}\}$ with $\text{Prob}\{\theta_i = \underline{\theta}\} = \text{v} > \frac{\Delta\theta}{\bar{\theta}}$ then $\mathbf{E}[\Pi^{FB} - \Pi^{SB}] = \frac{a^2}{2}\text{Prob}\{\theta_i = \underline{\theta}\}\mathcal{T}_i(\theta_i = \underline{\theta}, \bar{\theta}_{-i}) = \frac{a^3}{b^2}\,\text{v}\,\Delta\theta\,|\phi'(\bar{\theta})|\left[\sum_j k_i k_j g_{ij}\right]$, where the last equality follows by (36) and the fact that $\Delta\theta$ is sufficiently small. Therefore, it is immediate that $V^* = \arg\max_{i\in\{1,2,\cdots,n\}}\mathbf{E}[\Pi^{FB} - \Pi^{SB}] = \arg\max_{i\in\{1,2,\cdots,n\}}\sum_{j=1}^{n} k_i k_j g_{ij}$, completing the proof.

□

Proof of Proposition 3. We first derive the second-best trade profile. Define, for every i and $\theta_i \in [\underline{\theta}\ \bar{\theta}]$,

$$V_i(\theta_i) \triangleq ax_i(\theta_i) - \frac{b}{2}x_i^2(\theta_i) \tag{37}$$
$$\gamma_i(\theta_i) \triangleq x_i(\theta_i)\sum_{j\neq i} g_{ij}\mathbf{E}_{\theta_j}[x_j(\theta_j)] \tag{38}$$

Thus, the agent i's interim utility for reporting $\hat{\theta}_i$, while her real type is θ_i, is given by:

$$U_i(\theta_i, \hat{\theta}_i) = \theta_i\gamma_i(\hat{\theta}_i) + V_i(\hat{\theta}_i) - t_i(\hat{\theta}_i).$$

By appealing to the revenue equivalence theorem, a direct quantity-price schedule

$\{x_i^{SB}(\cdot), t_i^{SB}(\cdot)\}_{i=1}^n$ is IC and PC if and only if for every $i \in V$:

(i) $\gamma_i(\theta_i)$ is increasing;

(ii) For every $\theta_i \in [\underline{\theta}\ \bar{\theta}]$:

$$t_i(\theta_i) = \theta_i \gamma_i(\theta_i) + V_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau. \tag{39}$$

Given the specification of the incentive compatible and individual rational mechanisms the firm's problem is re-written as:

$$\begin{aligned} \max_{\{x_i(\cdot), t_i(\cdot)\}_{i=1}^n} \quad & \sum_{i=1}^n \mathbf{E}_{\theta_i}\Big[t_i(\theta_i)\Big] \\ \text{subject to} \quad & t_i(\theta_i) = \theta_i \gamma_i(\theta_i) + V_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau \quad \forall i,\ \theta_i, \qquad (40) \\ & \gamma_i(\theta_i) \text{ is increasing} \quad \forall\ i \qquad (41) \end{aligned}$$

To find the solution of the above program we first ignore the monotonicity constraints (41), later we show (41) is indeed satisfied. Note that $\mathbf{E}_{\theta_i}\left[\int_{\underline{\theta}}^{\theta_i} \gamma_i(\tau) d\tau\right] = \mathbf{E}_{\theta_i}\left[\frac{1-F(\theta_i)}{f(\theta_i)} \gamma_i(\theta_i)\right]$.

Hence, we obtain

$$\max_{\{x_i(\cdot)\}_{i=1}^n} \quad \sum_{i=1}^n \mathbf{E}_{\theta_i}\left[\psi(\theta_i)\gamma_i(\theta_i) + V_i(\theta_i)\right], \tag{42}$$

recall that $\psi(\theta_i) = \theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}$ (virtual type). Given the definitions in (37)-(38), and the fact that $\mathbf{E}[\cdot]$ is a linear operator, (42) can be re-written as:

$$\max_{\{x_i(\cdot)\}_{i=1}^n} \quad \mathbf{E} \sum_{i=1}^n \left[a x_i(\theta_i) - \frac{b}{2} x_i^2(\theta_i) + \psi(\theta_i) x_i(\theta_i) \sum_{j \neq i} g_{ij} \mathbf{E}[x_j(\tilde{\theta}_j)] \right] \tag{43}$$

To find the optimal solution to (43), we maximize it point-wise.

Note that $\sum_{i=1}^{n} \mathbf{E}\Big[\psi(\theta_i)x_i(\theta_i)\Big[\sum_{j\neq i} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\Big]$ can be rewritten as

$$
\begin{aligned}
&= \mathbf{E}\left[\psi_1(\theta_1)x_1(\theta_1)\Big[\sum_{j\neq 1} g_{1j}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\right] + \sum_{i=2}^{n}\mathbf{E}\left[\psi_i(\theta_i)x_i(\theta_i)\Big[\sum_{j\neq i} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\right] \\
&= \mathbf{E}\left[\psi_1(\theta_1)x_1(\theta_1)\Big[\sum_{j\neq 1} g_{1j}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\right] + \mathbf{E}\left[x_1(\theta_1)\Big[\sum_{j\neq 1} g_{j1}\mathbf{E}[\psi_j(\tilde{\theta}_j)x_j(\tilde{\theta}_j)]\Big]\right] \\
&\quad + \sum_{i=2}^{n}\mathbf{E}\left[\psi_i(\theta_i)x_i(\theta_i)\Big[\sum_{\substack{j\neq i\\ j\neq 1}} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\right] \\
&= \mathbf{E}\left[\psi_1(\theta_1)x_1(\theta_1)\Big[\sum_{j\neq 1} g_{1j}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big] + x_1(\theta_1)\Big[\sum_{j\neq 1} g_{j1}\mathbf{E}[\psi_j(\tilde{\theta}_j)x_j(\tilde{\theta}_j)]\Big]\right] \\
&\quad + \sum_{i=2}^{n}\mathbf{E}\left[\psi_i(\theta_i)x_i(\theta_i)\Big[\sum_{\substack{j\neq i\\ j\neq 1}} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big]\right] \qquad (44)
\end{aligned}
$$

Now, plugging (44) in the firm's problem, i.e. Eq. (43), and decoupling agent 1 from the rest, we obtain

$$
\begin{aligned}
&\sum_{i=1}^{n}\mathbf{E}\left[a x_i(\theta_i) - \frac{b}{2}x_i^2(\theta_i) + \psi(\theta_i)x_i(\theta_i)\sum_{j\neq i} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\right] \\
&= \sum_{i=2}^{n}\mathbf{E}\left[a x_i(\theta_i) - \frac{b}{2}x_i^2(\theta_i) + \psi(\theta_i)x_i(\theta_i)\sum_{\substack{j\neq i\\ j\neq 1}} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\right] \\
&+ \mathbf{E}\left[a_1 x_1(\theta_1) - \frac{b}{2}x_1^2(\theta_1) + \psi(\theta_1)x_1(\theta_1)\Big[\sum_{j\neq 1} g_{1j}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big] + x_1(\theta_1)\sum_{j\neq 1} g_{j1}\mathbf{E}[\psi_j(\tilde{\theta}_j)x_j(\tilde{\theta}_j)]\right].
\end{aligned}
$$

FOC with respect to $x_1(\theta)$, by keeping $\mathbf{x}_{-1}$ fixed, gives

$$
a - b x_1(\theta_1) + \psi(\theta_1)\Big[\sum_{j\neq 1} g_{1j}\mathbf{E}[x_j(\tilde{\theta}_j)]\Big] + \sum_{j\neq 1} g_{j1}\mathbf{E}[\psi_j(\tilde{\theta}_j)x_j(\tilde{\theta}_j)] = 0. \qquad (45)
$$

a similar equality can be obtained for any agent $i = 1, 2, \cdots, n$, that is for any i, we have

$$0 = a - bx_i(\theta_i) + \psi(\theta_i) \sum_{j=1}^{n} g_{ij} \mathbf{E}[x_j(\tilde{\theta}_j)] + \sum_{j=1}^{n} g_{ji} \mathbf{E}[\psi(\tilde{\theta}_j) x_j(\tilde{\theta}_j)]. \tag{46}$$

Equation (46) is rewritten in its matrix form as

$$b\mathbf{x}(\theta) = a\mathbf{1} + \psi(\theta)\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})]. \tag{47}$$

Taking expectation, we obtain

$$\begin{aligned} b\mathbf{E}[\mathbf{x}(\tilde{\theta})] &= a\mathbf{1} + \mathbf{E}[\psi(\tilde{\theta})]\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] \\ &= a\mathbf{1} + \mu_\psi \mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})]. \end{aligned} \tag{48}$$

Thus, we have

$$\mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] = (b\mathsf{I} - \mu_\psi \mathsf{G})\mathbf{E}[\mathbf{x}(\tilde{\theta})] - a\mathbf{1}. \tag{49}$$

Next, multiplying (47) by $\psi(\theta)$ and taking expectation[63], we obtain

$$\begin{aligned} b\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] &= \mathbf{E}[\psi(\tilde{\theta})]a\mathbf{1} + \mathbf{E}[\psi^2(\tilde{\theta})]\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathbf{E}[\psi(\tilde{\theta})]\mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] \\ &= \mu_\psi a\mathbf{1} + \mathbf{E}[\psi^2(\tilde{\theta})]\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mu_\psi \mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})]. \end{aligned} \tag{50}$$

Substituting for $\mathsf{G}^T \mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})]$ from (49), we have,

$$\begin{aligned} b\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] &= \mu_\psi a\mathbf{1} + \mathbf{E}[\psi^2(\tilde{\theta})]\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mu_\psi \Big((b\mathsf{I} - \mu_\psi \mathsf{G})\mathbf{E}[\mathbf{x}(\tilde{\theta})] - a\mathbf{1}\Big) \\ &= \mu_\psi b\ \mathbf{E}[\mathbf{x}(\tilde{\theta})] + \Big(\mathbf{E}[\psi^2(\tilde{\theta})] - \mu_\psi^2\Big)\mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] \\ &= \mu_\psi b\ \mathbf{E}[\mathbf{x}(\tilde{\theta})] + \sigma_\psi^2 \mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] \\ &= (\mu_\psi b\mathsf{I} + \sigma_\psi^2 \mathsf{G})\ \mathbf{E}[\mathbf{x}(\tilde{\theta})], \end{aligned}$$

where the second last line follows from the fact that $\sigma_\psi^2 = \mathbf{Var}[\psi(\tilde{\theta})] = \mathbf{E}[\psi^2(\tilde{\theta})] - \mu_\psi^2$.

[63] Note that since $\psi(\theta)$ and B are diagonal matrices we have, $\psi(\theta)B = B\psi(\theta)$.

Therefore,

$$\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] = \left(\mu_\psi \mathsf{I} + \frac{\sigma_\psi^2}{b}\mathsf{G}\right) \mathbf{E}[\mathbf{x}(\tilde{\theta})]. \tag{51}$$

Substituting the expression for $\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})]$ in (48), we obtain

$$\begin{aligned}
b\mathbf{E}[\mathbf{x}(\tilde{\theta})] &= a\mathbf{1} + \mu_\psi \mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathsf{G}^T\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] \\
&= a\mathbf{1} + \mu_\psi \mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] + \mathsf{G}^T\left(\mu_\psi \mathsf{I} + \frac{\sigma_\psi^2}{b}\mathsf{G}\right)\mathbf{E}[\mathbf{x}(\tilde{\theta})] \\
&= a\mathbf{1} + \left[\mu_\psi(\mathsf{G}+\mathsf{G}^T) + \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]\mathbf{E}[\mathbf{x}(\tilde{\theta})].
\end{aligned} \tag{52}$$

Therefore, we obtain the following linear systems of equations:

$$\left[b\mathsf{I} - \mu_\psi(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]\mathbf{E}[\mathbf{x}(\tilde{\theta})] = a\mathbf{1}.$$

From Assumption 3, we know that the coefficient matrix is invertible. Thus, we obtain,

$$\mathbf{E}[\mathbf{x}(\tilde{\theta})] = \left[b\mathsf{I} - \mu_\psi(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1} a\mathbf{1} = a\mathsf{K}_\psi\mathbf{1}, \tag{53}$$

where $\mathsf{K}_\psi \triangleq \left[b\mathsf{I} - \mu_\psi(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1}$.

Finally, from equations (47), (49), and (53), we obtain,

$$\begin{aligned}
\mathbf{x}(\theta) &= b^{-1}\left[a\mathbf{1} + \psi(\theta)\mathsf{G}\mathsf{K}_\psi\Lambda + \left((b\mathsf{I} - \mu_\psi\mathsf{G})\mathsf{K}_\psi a\mathbf{1} - a\mathbf{1}\right)\right] \\
&= b^{-1}\left[\psi(\theta)\mathsf{G}\mathsf{K}_\psi a\mathbf{1} + \left(b\mathsf{I} - \mu_\psi\mathsf{G}\right)\mathsf{K}_\psi a\mathbf{1}\right] \\
&= b^{-1}\left[(\psi(\theta) - \mu_\psi)\mathsf{G}\mathsf{K}_\psi a\mathbf{1} + b\mathsf{K}_\psi a\mathbf{1}\right] \\
&= b^{-1}\left(\psi(\theta) - \mu_\psi\right)\mathsf{G}\mathsf{K}_\psi a\mathbf{1} + \mathsf{K}_\psi a\mathbf{1}.
\end{aligned} \tag{54}$$

Therefore,

$$x^{SB}(\theta_i) = \frac{a}{b}(\psi(\theta_i) - \mu_\psi)[\mathsf{G}\mathsf{K}_\psi \mathbf{1}]_i + a[\mathsf{K}_\psi \mathbf{1}]_i.$$

To complete the proof, we next show $\underline{\mathbf{x}}^{SB} > \mathbf{0}$. Note that $\mu_\psi = \mathbf{E}[\psi(\theta_i)] = \underline{\theta}$. Thus,

$$\psi(\underline{\theta}) - \mu_\psi = -\frac{1 - F(\underline{\theta})}{f(\underline{\theta})} = -\frac{1}{f(\underline{\theta})}.$$

Define, $\alpha \triangleq \frac{1}{b}\,\frac{1}{f(\underline{\theta})}$. By Assumption 3, $b\mathsf{I} - \mu_\psi(\mathsf{G} + \mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}$ is invertible. Thus,

$$\begin{aligned}
\mathsf{K}_\psi a\mathbf{1} &= \left[b\mathsf{I} - \mu_\psi(\mathsf{G} + \mathsf{G}^T) - \frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1} a\mathbf{1} \\
&= \left[\mathsf{I} - \frac{\mu_\psi}{b}(\mathsf{G} + \mathsf{G}^T) - \frac{\sigma_\psi^2}{b^2}\mathsf{G}^T\mathsf{G}\right]^{-1} \frac{a}{b}\mathbf{1} \\
&= (I - A)^{-1}\frac{a}{b}\mathbf{1},
\end{aligned}$$

where $A \triangleq \frac{\mu_\psi}{b}(\mathsf{G} + \mathsf{G}^T) + \frac{\sigma_\psi^2}{b^2}\mathsf{G}^T\mathsf{G}$. Then, we have for $\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2, \cdots, \underline{\theta}_n)$:

$$\begin{aligned}
\mathbf{x}(\theta) &\geq \mathbf{x}(\underline{\theta}) \\
&= (\mathsf{I} - \alpha\mathsf{G})\mathsf{K}_\psi a\mathbf{1} \\
&= (\mathsf{I} - \alpha\mathsf{G})(\mathsf{I} - A)^{-1} b^{-1} a\mathbf{1} \\
&\overset{(a)}{=} (\mathsf{I} - \alpha\mathsf{G})(\mathsf{I} + A + A^2 + A^3 + \cdots)b^{-1}a\mathbf{1} \\
&= \Big(\mathsf{I} - \alpha\mathsf{G} + A - \alpha\mathsf{G}A + A^2 - \alpha\mathsf{G}A^2 + A^3 + \cdots\Big)b^{-1}a\mathbf{1} \\
&= \Big(\mathsf{I} + (A - \alpha\mathsf{G}) + (A - \alpha\mathsf{G})A + (A - \alpha\mathsf{G})A^2 + \cdots\Big)b^{-1}a\mathbf{1} \\
&> 0,
\end{aligned}$$

where (a) follows since by Assumption 3, $\rho(A) < 1$, thus $(\mathsf{I} - A)^{-1} = I + A + A^2 + \cdots$, and the last inequality is correct because $b^{-1}a > 0$ and $A - \alpha\mathsf{G} \geq \mathbf{0}$ (since $\psi(\underline{\theta}) \geq 0$). Finally, monotonicity of $x^{SB}(\theta_i)$ is immediate since $\psi(\theta_i)$ is monotone in θ_i, completing the proof.

Full information (first-best): In this case there is no IC and all the PC constraint must bind, characterizing the payments. Thus:

$$t_i(\theta_i) = ax_i(\theta_i) - \frac{b}{2}x_i^2(\theta_i) + \theta_i x_i(\theta_i)\sum_{j=1}^{n} g_{ij}\mathbf{E}[x_j(\theta_j)].$$

Hence, the firms problem is written as

$$\max_{x,t} \quad \sum_{i=1}^{n}\mathbf{E}[t_i(\theta_i)]$$
$$\text{subject to} \qquad \text{PC.}$$

Plugging the payments in the objective function we have, and using FOC with respect to $x_i(\theta_i)$ give

$$0 = a - bx_i(\theta_i) + \theta_i\sum_{j=1}^{n} g_{ij}\mathbf{E}[x_j(\theta_j)] + \sum_{j=1}^{n} g_{ji}\mathbf{E}[\theta_j x_j(\theta_j)]. \tag{55}$$

Eqs. (55) can be re-written in the following compact form:

$$0 = a - bx_i(\theta_i) + \theta_i\sum_{j=1}^{n} g_{ij}\mathbf{E}[x_j(\theta_j)] + \sum_{j=1}^{n} g_{ji}\mathbf{E}[\theta_j x_j(\theta_j)], \tag{56}$$

where $\theta_i \in [\underline{\theta}, \bar{\theta}]$, for all i. The above equality is similar to (46), thus, by following the same argument as in the second-best we obtain:

$$x_i^{FB}(\theta_i) = \frac{a}{b}(\theta_i - \mu)[\mathsf{GK}\mathbf{1}]_i + a[\mathsf{K}\mathbf{1}]_i,$$

where $\mathsf{K} \triangleq \left[b\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \frac{\sigma^2}{b}\mathsf{G}^T\mathsf{G}\right]^{-1}$. □

Proof of Proposition 4. The first part is immediate from Proposition 3. For the second, consider the full information case. Using Proposition 3, $x_i^{FB}(\theta_i) = \frac{a}{b}(\theta_i - \mu)[\mathsf{GK}\mathbf{1}]_i + a[\mathsf{K}\mathbf{1}]_i$ where $\mathsf{K} \triangleq \left[b\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \frac{1}{b}\mathsf{G}^T\Sigma\mathsf{G}\right]^{-1}$, and Σ is a diagonal matrix where $Var[\theta_i] = \sigma_i^2$ is in entry ii. Since $\mathsf{I} = \mathsf{K}\mathsf{K}^{-1}$, then by chain rule we obtain,

$$\mathbf{0} = \frac{\partial \mathsf{I}}{\partial \sigma_i^2} = \frac{\partial \mathsf{K}\mathsf{K}^{-1}}{\partial \sigma_i^2} = \frac{\partial \mathsf{K}}{\partial \sigma_i^2}\mathsf{K}^{-1} + \mathsf{K}\frac{\partial \mathsf{K}^{-1}}{\partial \sigma_i^2}. \tag{57}$$

where $\mathbf{0}$ denotes a $n \times n$ zero matrix. Then

$$\frac{\partial \mathsf{K}^{-1}}{\partial \sigma_i^2} = \frac{\partial}{\partial \sigma_i^2}\left[b\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \frac{1}{b}\mathsf{G}^T\Sigma\mathsf{G}\right] = -\frac{1}{b}\frac{\partial}{\partial \sigma_i^2}\left[\mathsf{G}^T\Sigma\mathsf{G}\right] = -\frac{1}{b}\mathsf{G}^T\ \frac{\partial \Sigma}{\partial \sigma_i^2}\ \mathsf{G} = -\frac{1}{b}\mathsf{G}^T E_{ii}\ \mathsf{G},$$

where $E_{ii} = \frac{\partial \Sigma}{\partial \sigma_i^2}$ is a matrix containing only one non-zero entry equal to one located at the intersection of the i'th row and i'th column. Finally, using the last equality in (57) we obtain

$$\frac{\partial \mathsf{K}}{\partial \sigma_i^2} = -\mathsf{K}\ \frac{\partial \mathsf{K}^{-1}}{\partial \sigma_i^2}\ \mathsf{K} = \frac{1}{b}\mathsf{K}\mathsf{G}^T E_{ii}\mathsf{G}\mathsf{K}.$$

From the last equality, since all the matrices are component-wise non-negative, we imply that $\frac{\partial \mathsf{K}}{\partial \sigma_i^2}$ is a matrix with non-negative entries, that completes the proof. We note that one way to change σ_i while keeping its mean fixed is to change the support of $\Theta_i = [\underline{\theta}\ \bar{\theta}]$. The proof for the imperfect information case is similar. □

Proof of Proposition 5. Given (14), using $\mathbf{E}[\psi(\tilde{\theta}) - \mu_\psi] = 0$, we have

$$\eta \triangleq \mathsf{G}\mathbf{E}[\mathbf{x}(\tilde{\theta})] = \mathsf{G}\mathsf{K}_\psi\Lambda, \tag{58}$$

where $\eta = (\eta_1, \eta_2, \cdots, \eta_n)^T$ and $\Lambda \triangleq a\mathbf{1}$.

In order to derive $\mathbf{E}\left[\sum_{i=1}^n t_i(\theta_i)\right]$, we decompose it into three terms as follows

$$\begin{aligned}\mathbf{E}\left[\sum_{i=1}^n t_i(\theta_i)\right] &= \mathbf{E}\sum_{i=1}^n\left[a x_i(\theta_i) - \frac{b}{2}x_i^2(\theta_i) + \psi(\theta_i)x_i(\theta_i)\sum_{j\neq i} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\right] \\ &= \zeta_1 + \zeta_2 + \zeta_3,\end{aligned}$$

where

$$\zeta_1 = \sum_{i=1}^{n} \mathbf{E}[\psi(\theta_i)x_i(\theta_i)]\left[\sum_{j\neq i} g_{ij}\mathbf{E}[x_j(\tilde{\theta}_j)]\right] = \sum_{i=1}^{n} \eta_i \mathbf{E}[\psi(\theta_i)x_i(\theta_i)],$$

$$\zeta_2 = \frac{-b}{2}\sum_{i=1}^{n} \mathbf{E}[x_i^2(\theta_i)],$$

$$\zeta_3 = a\sum_{i=1}^{n} \mathbf{E}[x_i(\theta_i)].$$

It follows that

$$\begin{aligned}\zeta_1 &= \sum_{i=1}^{n} \eta_i\ \mathbf{E}[\psi_i(\tilde{\theta}_i)x_i(\tilde{\theta}_i)] = \eta^T\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] \overset{(a)}{=} (\mathsf{G}\mathsf{K}_\psi\Lambda)^T\left(\mu_\psi \mathsf{I} + \frac{\sigma_\psi^2}{b}\mathsf{G}\right)\mathsf{K}_\psi\Lambda \\ &= \frac{\sigma_\psi^2}{b}\Lambda^T\mathsf{K}_\psi^T\mathsf{G}^T\mathsf{G}\mathsf{K}_\psi\Lambda + \mu_\psi\Lambda^T\mathsf{K}_\psi^T\mathsf{G}^T\mathsf{K}_\psi\Lambda, \end{aligned} \tag{59}$$

where (a) follows from (51), (that is, $\mathbf{E}[\psi(\tilde{\theta})\mathbf{x}(\tilde{\theta})] = \left(\mu_\psi + \frac{\sigma_\psi^2}{b}\mathsf{G}\right)\mathbf{E}[\mathbf{x}(\tilde{\theta})] = \left(\mu_\psi + \frac{\sigma_\psi^2}{b}\mathsf{G}\right)\mathsf{K}_\psi\Lambda$).

Furthermore, we obtain

$$\begin{aligned}\zeta_2 &= \frac{-b}{2}\mathbf{E}\left[\sum_{i=1}^{n} x_i^2(\tilde{\theta}_i)\right] = \frac{-b}{2}\mathbf{E}\left[\mathbf{x}(\tilde{\theta})^T\ \mathbf{x}(\tilde{\theta})\right] \\ &= \frac{-b}{2}\mathbf{E}\left[\left[\frac{1}{b}(\psi(\tilde{\theta}) - \mu_\psi)\mathsf{G}\mathsf{K}_\psi\Lambda + \mathsf{K}_\psi\Lambda\right]^T\left[\frac{1}{b}(\psi(\tilde{\theta}) - \mu_\psi)\mathsf{G}\mathsf{K}_\psi\Lambda + \mathsf{K}_\psi\Lambda\right]\right] \\ &= \frac{-b}{2}\mathbf{E}\left[\left[\Lambda^T\mathsf{K}_\psi^T\mathsf{G}^T(\psi(\tilde{\theta}) - \mu_\psi)\frac{1}{b} + \Lambda^T\mathsf{K}_\psi^T\right]\left[\frac{1}{b}(\psi(\tilde{\theta}) - \mu_\psi)\mathsf{G}\mathsf{K}_\psi\Lambda + \mathsf{K}_\psi\Lambda\right]\right] \\ &\overset{(a)}{=} \frac{-\sigma_\psi^2}{2b}\Lambda^T\mathsf{K}_\psi^T\mathsf{G}^T\mathsf{G}\mathsf{K}_\psi\Lambda - \frac{b}{2}\Lambda^T\mathsf{K}_\psi^T\mathsf{K}_\psi\Lambda, \end{aligned} \tag{60}$$

where (a) is followed by the fact that $\mathbf{E}[(\psi(\tilde{\theta}) - \mu_\psi)] = 0$ and $\mathbf{E}[(\psi(\tilde{\theta}) - \mu_\psi)^2] = \sigma_\psi^2$.

Following the same arguments we also obtain

$$\zeta_3 = a\mathbf{E}[\mathbf{x}(\tilde{\theta})] = a\mathbf{1}^T\mathsf{K}_\psi\Lambda = \Lambda^T\mathsf{K}_\psi\Lambda. \tag{61}$$

Now, from (59), (60) and (61) we obtain

$$\Pi_{\mathsf{G}} = \mathbf{E}\left[\sum_{i=1}^{n} t_i(\theta_i)\right] = \zeta_1 + \zeta_2 + \zeta_3$$
$$= \frac{\sigma_\psi^2}{2b}\Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{G}\mathsf{K}_\psi\Lambda + \mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda - \frac{b}{2}\Lambda^T \mathsf{K}_\psi^T \mathsf{K}_\psi \Lambda + \Lambda^T \mathsf{K}_\psi \Lambda. \quad (62)$$

Note that, since $\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda$ is a scalar, thus

$$\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda = \left(\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda\right)^T = \mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}\mathsf{K}_\psi \Lambda.$$

Thus

$$\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda = \frac{1}{2}\left[\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda + \left(\mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda\right)^T\right]$$
$$= \frac{1}{2}\left[\Lambda^T \mathsf{K}_\psi^T \left[\mu_\psi\left(\mathsf{G}^T + \mathsf{G}\right)\right] \mathsf{K}_\psi \Lambda\right].$$

Hence by the preceding equality, we can simplify (62) as

$$\begin{aligned}
\Pi_{\mathsf{G}} &= \frac{\sigma_\psi^2}{2b}\Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{G}\mathsf{K}_\psi\Lambda + \mu_\psi \Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{K}_\psi \Lambda - \frac{b}{2}\Lambda^T \mathsf{K}_\psi^T \mathsf{K}_\psi \Lambda + a\mathsf{K}_\psi \Lambda \\
&= \frac{\sigma_\psi^2}{2b}\Lambda^T \mathsf{K}_\psi^T \mathsf{G}^T \mathsf{G}\mathsf{K}_\psi\Lambda + \frac{1}{2}\left[\Lambda^T \mathsf{K}_\psi^T \left[\mu_\psi\left(\mathsf{G}^T + \mathsf{G}\right)\right] \mathsf{K}_\psi \Lambda\right] - \frac{b}{2}\Lambda^T \mathsf{K}_\psi^T \mathsf{K}_\psi \Lambda + \Lambda^T \mathsf{K}_\psi \Lambda \\
&= \frac{1}{2}\Lambda^T \mathsf{K}_\psi^T \left[\frac{\sigma_\psi^2}{b}\mathsf{G}^T\mathsf{G} + \mu_\psi(\mathsf{G}^T + \mathsf{G}) - b\mathsf{I}\right]\mathsf{K}_\psi \Lambda + \Lambda^T \mathsf{K}_\psi \Lambda \\
&\overset{a}{=} -\frac{1}{2}\Lambda^T \mathsf{K}_\psi^T \mathsf{K}_\psi^{-1} \mathsf{K}_\psi \Lambda + \Lambda^T \mathsf{K}_\psi \Lambda \\
&= -\frac{1}{2}\Lambda^T \mathsf{K}_\psi^T \Lambda + \Lambda^T \mathsf{K}_\psi \Lambda \\
&\overset{b}{=} \frac{1}{2}\Lambda^T \mathsf{K}_\psi \Lambda \\
&\overset{c}{=} \frac{a^2}{2}\, \mathbf{1}^T \mathsf{K}_\psi \mathbf{1},
\end{aligned}$$

where (a) follows from the definition of K_ψ and (b) follows from symmetry of K_ψ, i.e., $\mathsf{K}_\psi = \mathsf{K}_\psi^T$ and (c) follows because $\Lambda = a\mathbf{1}$ by its definition. The proof for the full information case follows similar steps and the final answer becomes $\Pi_{\mathsf{G}}^{bi.} = \frac{a^2}{2}\, \mathbf{1}^T \mathsf{K}\mathbf{1}$.

□

Proof of Proposition 6. Recall that $b = 1$ and we are in the full information case. The same proof works for the incomplete information case as well.

Following the same steps as in the proof of Proposition 5, one can show

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\,\mathbf{1}^T \mathsf{K}\mathbf{1} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \sigma^2\mathsf{G}^T\mathsf{G}\right]^{-1}\mathbf{1}.$$

Since σ^2 is chosen to be small, thus Taylor expansion of $\Pi_{\mathsf{G}}^{\text{bi.}}$ around $\sigma^2 = 0$ gives

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1} + \sigma^2\,\mathsf{S}^{-1}\,\mathsf{G}^T\mathsf{G}\,\mathsf{S}^{-1}\right]\mathbf{1},$$

where $\mathsf{S}^{-1} \triangleq [\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T)]^{-1}$.

Since G is symmetric (i.e. $\mathsf{G} = \mathsf{G}^T$) it has n district eigenvalues and is thus diagonalizable. Therefore, there exists an invertible matrix V so that $V^{-1}\mathsf{G}V = \Lambda$, where Λ is a diagonal matrix with the eigenvalues of G on its diagonal, and columns of V are the eigenvectors of G. Therefore, $\mathsf{G} = V\Lambda V^{-1}$. In addition, the eigenvectors can be chosen to form an orthonormal basis of $\mathbb{R}^n$, meaning that V is set to be orthonormal. That is $V^TV = VV^T = V^{-1}V = \mathsf{I}$. Moreover,

$$\mathsf{S}^{-1} = \sum_{i=0}^{\infty}(2\mu\,\mathsf{G})^i = \sum_{i=0}^{\infty}\left(2\mu\,V\Lambda V^{-1}\right)^i = V\left(\sum_{i=0}^{\infty}(2\mu\,\Lambda)^i\right)V^{-1} = V\Lambda_1 V^{-1}, \tag{63}$$

where Λ_1 is a diagonal matrix where its k–th element is $\frac{1}{1-2\mu\lambda_k}$, (note that (by Assumption 3) $2\mu\lambda_k \leq 2\mu\lambda_{max} < 1$, for all k). By following similar argument as in (63), one can also show:

$$\mathsf{S}^{-1}\,\mathsf{G}^T\mathsf{G}\,\mathsf{S}^{-1} = V\Lambda_1^2\Lambda^2 V^{-1}. \tag{64}$$

Eq. (63) together with (64) yield:

$$\Xi \triangleq \mathsf{S}^{-1} + \sigma^2\,\mathsf{S}^{-1}\,\mathsf{G}^T\mathsf{G}\,\mathsf{S}^{-1} = V\left(\underbrace{\Lambda_1 + \sigma^2(\Lambda_1\Lambda)^2}_{\text{diagonal}}\right)V^{-1}. \tag{65}$$

The above equality characterizes eigenvalues of Ξ in terms of $\lambda_1, \lambda_2, \cdots, \lambda_n$. That is, k–th eigenvalue of Ξ is equal to $f(\lambda_k) = \frac{1-2\mu\lambda_k+\sigma^2\lambda_k^2}{(1-2\mu\lambda_k)^2}$. Further, it can be easily shown $f(\lambda)$ is increasing and convex in λ.

To wrap up the proof we make a use of the following Lemma.

Lemma 6. *Let ζ_{min} and ζ_{max} be the smallest and largest eigenvalues of the square matrix M (with n distinct eigenvalues). Then:*

$$\zeta_{min} x^T x \leq x^T M x \leq \zeta_{max} x^T x,$$

where $x \in \mathbb{R}^n$.

Proof. Let $v_1, v_2, \cdots, v_n$ denote the eigenvectors corresponding to the eigenvalues $\zeta_1, \zeta_2, \cdots, \zeta_n$ of M. Thus, there exists $\alpha_1, \cdots, \alpha_n$ such that $x = \sum_{i=1}^n \alpha_i v_i$. Using this fact, since

$$Mx = M\sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \alpha_i \; M v_i = \sum_{i=1}^n \alpha_i \; \zeta_i v_i \leq \zeta_{max} \sum_{i=1}^n \alpha_i v_i = \zeta_{max} x,$$

thus $x^T M x \leq \zeta_{max} x^T x$. And similarly the lower can be proved, completing the proof. □

Finally, the proof of the proposition is immediate by employing Lemma 6 and the fact that $\mathbf{1}^T\mathbf{1} = n$, i.e.,

$$0 < \frac{na^2}{2} f(\lambda_{min}) \leq \Pi_{\mathsf{G}}^{bi.} \leq \frac{na^2}{2} f(\lambda_{\max})$$

where $f(\lambda) = \frac{1-2\mu\lambda+\sigma^2\lambda^2}{(1-2\mu\lambda)^2}$.

□

Proof of Proposition 15. The proof follows similar steps as in the proof of Proposition 4 and is omitted. □

Proof of Corollary 2. Given Proposition 5, the proof is immediate because

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2} \mathbf{1}^T \mathsf{K} \mathbf{1} = \frac{a^2}{2} \mathbf{1}^T \left[\mathsf{I} - \mu(\mathsf{G} + \mathsf{G}^T) - \sigma^2 \mathsf{G}^T \mathsf{G} \right]^{-1} \mathbf{1}$$

thus $\frac{\partial \Pi_{\mathsf{G}}^{bi.}}{\partial \sigma^2} = \frac{a^2}{2b} \mathbf{1}^T \mathsf{K} \, \mathsf{G}^T \mathsf{G} \, \mathsf{K} \mathbf{1} > 0$. The proof for the imperfect information case is similar.

To gain intution, in the following we consider a discrete distribution and change σ^2 while the mean is kept fixed and prove the corollary. Let us assume $\text{Prob}\{\theta_i = \underline{\theta}\} = 1 - \text{Prob}\{\theta_i = \bar{\theta}\} = \mathsf{v} > \frac{\Delta\theta}{\bar{\theta}}$, for all i. Without loss of generality, let $\bar{\theta} = \delta\,\underline{\theta} > 0$, for some δ. Further, since $\mathsf{v} > \frac{\Delta\theta}{\bar{\theta}}$, thus $\delta > \frac{1}{1-\mathsf{v}}$. Since $\mu = \mathbf{E}[\theta_i]$ is fixed and given, thus $\underline{\theta} = \frac{\mu}{1+(1-\mathsf{v})(\delta-1)}$. As a result,

$$\sigma^2 = \mathsf{v}(1-\mathsf{v})(\bar{\theta}-\underline{\theta})^2 = \mathsf{v}(1-\mathsf{v})\mu^2\left(\frac{\delta-1}{1+(1-\mathsf{v})(\delta-1)}\right)^2.$$

We note that since μ and v are fixed and given, thus σ^2 is only controlled by δ. It also observes that σ^2 is increasing in δ, i.e. $\frac{\partial\sigma^2}{\partial\delta} > 0$. Intuitively, increasing δ, increases $\Delta\theta = \bar{\theta} - \underline{\theta}$, while μ and v are both kept fixed, increasing σ^2. Thus, given Proposition 5, $\frac{\partial\Pi_{\mathsf{G}}^{bi.}}{\partial\delta} = \frac{\partial}{\partial\sigma^2}[\frac{a^2}{2}\,\mathbf{1}^T\mathsf{K}\mathbf{1}]\frac{\partial\sigma^2}{\partial\delta} = [\frac{a^2}{2}\,\mathbf{1}^T\frac{\partial\mathsf{K}}{\partial\sigma^2}\mathbf{1}]\frac{\partial\sigma^2}{\partial\delta} = [\frac{a^2}{2b}\,\mathbf{1}^T\mathsf{K}\,\mathsf{G}^T\mathsf{G}\,\mathsf{K}\mathbf{1}]\frac{\partial\sigma^2}{\partial\delta} > 0.$ □

Proof of Proposition 7. Recall that $b = 1$ and we are in the full information case. For ease of exposition define $\mathsf{S}^{-1} \triangleq (\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T))^{-1}$. Following the same steps as in the proof of Proposition 5, one can show

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\,\mathbf{1}^T\mathsf{K}\mathbf{1} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \sigma^2\mathsf{G}^T\mathsf{G}\right]^{-1}\mathbf{1}.$$

Since σ^2 is chosen to be small, thus Taylor expansion of $\Pi_{\mathsf{G}}^{\text{bi.}}$ around $\sigma^2 = 0$ implies

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1} + \sigma^2\,\mathsf{S}^{-1}\,\mathsf{G}^T\mathsf{G}\,\mathsf{S}^{-1}\right]\mathbf{1}.$$

Since G is balanced (see Def. 4), thus $\mathsf{S}^{-1}\mathbf{1} = \frac{1}{1-\mu\tau}\mathbf{1} = \zeta\mathbf{1}$ (note that τ is small enough so that all the matrix-invertibilities are preserved). Therefore, $\mathbf{1}^T\mathsf{S}^{-1}\mathbf{1} = \zeta n$ and $\mathbf{1}^T\mathsf{S}^{-1}\mathsf{G}^T\mathsf{G}\mathsf{S}^{-1}\mathbf{1} = \zeta^2\mathbf{1}^T\mathsf{G}^T\mathsf{G}\mathbf{1} = \sum_i\left(\sum_j g_{ij}\right)^2$, completing the proof.

The same argument holds for the second-best contract, with the change that linearization (Taylor expansion) of $\Pi_{\mathsf{G}}^{\text{bi.,SB}}$ is around $\sigma_\psi^2 = 0$.

□

Proof of Lemma 2. Recall that $b = 1$ and we are in the first-best contract. The same proof works for the second-best contract as well, note that there only θ changes to $\psi(\theta)$ and Taylor expansion will be around σ_ψ^2.

Following the same steps as in the proof of Proposition 5, one can show

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\,\mathbf{1}^T\mathsf{K}\mathbf{1} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{I} - \mu(\mathsf{G}+\mathsf{G}^T) - \sigma^2\mathsf{G}^T\mathsf{G}\right]^{-1}\mathbf{1}.$$

Since σ^2 is chosen to be small, thus Taylor expansion of $\Pi_{\mathsf{G}}^{\text{bi.}}$ around $\sigma^2 = 0$ gives

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1} + \sigma^2\,\mathsf{S}^{-1}\,\mathsf{G}^T\mathsf{G}\,\mathsf{S}^{-1}\right]\mathbf{1}.$$

By little algebra, it can be easily shown that

$$\mathsf{G}^T\mathsf{G} = \sum_i R_i^2 - \operatorname{diag}\left(\sum_{j=1}^n g_{1j}^2, \sum_{j=1}^n g_{2j}^2, \cdots, \sum_{j=1}^n g_{nj}^2\right)$$

that implies:

$$\Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1} + \sigma^2\,\mathsf{S}^{-1}\,\left(\sum_i R_i^2 - \operatorname{diag}\left(\sum_{j=1}^n g_{1j}^2, \sum_{j=1}^n g_{2j}^2, \cdots, \sum_{j=1}^n g_{nj}^2\right)\right)\,\mathsf{S}^{-1}\right]\mathbf{1}. \tag{66}$$

Next, we proceed with computing $\Pi_{\mathsf{G}}^{\text{multi.}}$. As shown in Proposition 1, the optimal (first-best) trade profile is $\mathbf{x}^{FB}(\theta) = a\,[b\mathsf{I} - (\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1}\mathbf{1}$. Thus, for each type profile θ the (ex post) firm's profit is given by (let x_i denote agent i's allocation with respect to the type profile θ)

$$\begin{aligned}
\sum_i\left(ax_i - \frac{b}{2}x_i^2 + x_i\theta_i\sum_j g_{ij}x_j\right) &= a\mathbf{1}^T\mathbf{x} - \frac{b}{2}\mathbf{x}^T\mathbf{x} + (\mathsf{M}_\theta\mathbf{x})^T\mathsf{G}\mathbf{x} \\
&= a\mathbf{1}^T\mathbf{x} - \frac{b}{2}\mathbf{x}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T(\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)\mathbf{x} \\
&= a\mathbf{1}^T\mathbf{x} - \frac{1}{2}\left[\mathbf{x}^T\;\;[b\mathsf{I} - (\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]\;\;\mathbf{x}\right] \\
&= \frac{a^2}{2}\mathbf{1}^T[b\mathsf{I} - (\mathsf{M}_\theta\mathsf{G} + \mathsf{G}^T\mathsf{M}_\theta)]^{-1}\mathbf{1},
\end{aligned}$$

where the last equality follows by substituting the optimal allocation trade in it. Therefore, (setting $b=1$) the (ex-ante) firm's profit (using multilateral) contract becomes:

$$\Pi_{\mathsf{G}}^{\text{multi.}} = \frac{a^2}{2}\mathbf{1}^T\ \mathbf{E}\left[\mathsf{I}-(\mathsf{M}_\theta\mathsf{G}+\mathsf{G}^T\mathsf{M}_\theta)\right]^{-1}\ \mathbf{1}. \tag{67}$$

To simplify $\mathbf{E}\left[\mathsf{I}-(\mathsf{M}_\theta\mathsf{G}+\mathsf{G}^T\mathsf{M}_\theta)\right]^{-1}$, we first write the Taylor expansion of $\Lambda \triangleq \left[\mathsf{I}-(\mathsf{M}_\theta\mathsf{G}+\mathsf{G}^T\mathsf{M}_\theta)\right]^{-1}$ around $\theta=\mu\mathbf{1}$. Thus we have:

$$\begin{aligned}\left[\mathsf{I}-(\mathsf{M}_\theta\mathsf{G}+\mathsf{G}^T\mathsf{M}_\theta)\right]^{-1} &= \left[\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T)\right]^{-1} + \sum_i(\theta_i-\mu)\left(\frac{\partial\Lambda}{\partial\theta_i}|_{\theta=\mu\mathbf{1}}\right)\\ &+\frac{1}{2!}\sum_i\sum_j(\theta_i-\mu)(\theta_j-\mu)\left(\frac{\partial^2\Lambda}{\partial\theta_i\partial\theta_j}|_{\theta=\mu\mathbf{1}}\right)\\ &+\sum_{n=3}^{\infty}\frac{1}{n!}\sum_{\substack{\zeta_1,\zeta_2,\cdots,\zeta_n\geq 0\\ \zeta_1+\cdots+\zeta_n=n}}\left[\prod_{k=1}^{n}(\theta_k-\mu)^{\zeta_k}\right]\left(\frac{\partial^n\Lambda}{\partial\theta_1^{\zeta_1}\partial\theta_2^{\zeta_2}\cdots\partial\theta_n^{\zeta_n}}|_{\theta=\mu\mathbf{1}}\right).\end{aligned} \tag{68}$$

Next, since $\mathbf{E}(\theta_i-\mu)=0$, for all i, we have

$$\begin{aligned}\mathbf{E}\left[\left[\mathsf{I}-(\mathsf{M}_\theta\mathsf{G}+\mathsf{G}^T\mathsf{M}_\theta)\right]^{-1}\right] &= \left[\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T)\right]^{-1}+\frac{1}{2!}\sum_i\mathbf{E}[(\theta_i-\mu)^2]\left(\frac{\partial^2\Lambda}{\partial\theta_i\partial\theta_j}|_{\theta=\mu\mathbf{1}}\right)\\ &+T_{>2},\end{aligned} \tag{69}$$

where

$$T_{>2}=\sum_{n=3}^{\infty}\frac{1}{n!}\sum_{\substack{\zeta_1,\zeta_2,\cdots,\zeta_n\\ \zeta_1+\cdots+\zeta_n=n\\ \zeta_i\neq 1,\forall i}}\prod_{k=1}^{n}\mathbf{E}[(\theta_k-\mu)^{\zeta_k}]\left(\frac{\partial^n\Lambda}{\partial\theta_1^{\zeta_1}\partial\theta_2^{\zeta_2}\cdots\partial\theta_n^{\zeta_n}}|_{\theta=\mu\mathbf{1}}\right).$$

Notice that in $T_{>2}, \zeta_i\neq 1$ (for all i), since $\mathbf{E}(\theta_i-\mu)=0$. Moreover, by the Assumption, since there exists $\hat{m}>0$, such that $\mathbf{E}[(\theta_i-\mu)^k]<(\hat{m}\sigma)^k$, for all i and $k\geq 3$, thus for all $n\geq 3$,

$$\prod_{k=1}^{n}\mathbf{E}[(\theta_k-\mu)^{\zeta_k}]<(\hat{m}\sigma)^{\sum_k\zeta_k}=(\hat{m}\sigma)^n.$$

In addition, σ^2 is chosen to be small, implying $\hat{m}\sigma$ is small (since $\hat{m}$ is constant). Consequently,

$$T_{>2} < \sum_{n=3}^{\infty} \frac{1}{n!} \sum_{\substack{\zeta_1,\zeta_2,\cdots,\zeta_n \\ \zeta_1+\cdots+\zeta_n=n \\ \zeta_i \neq 1, \forall i}} (\hat{m}\sigma)^n \left(\frac{\partial^n \Lambda}{\partial\theta_1^{\zeta_1} \partial\theta_2^{\zeta_2} \cdots \partial\theta_n^{\zeta_n}} |_{\theta=\mu\mathbf{1}} \right).$$

Above inequality along with the Taylor expansion in (68) and the fact that $\hat{m}$ is constant imply that

$$T_{>2} = O(\sigma^3).$$

Therefore, since σ^2 is chosen to be small, $T_{>2}$ is negligible and thus (69) becomes

$$\mathbf{E}\left[\left[\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta) \right]^{-1} \right] = \left[\mathsf{I} - \mu(\mathsf{G} + \mathsf{G}^T) \right]^{-1} + \frac{\sigma^2}{2!} \sum_i \left(\frac{\partial^2 \Lambda}{\partial \theta_i^2} |_{\theta=\mu\mathbf{1}} \right) \tag{70}$$

Moreover, for all i:

$$\frac{\partial^2 \Lambda}{\partial \theta_i^2} |_{\theta=\mu\mathbf{1}} = 2\mathsf{S}^{-1} \left(R_i \, \mathsf{S}^{-1} \, R_i \right) \mathsf{S}^{-1}.$$

Thus, (70) is simplified as

$$\mathbf{E}\left[\left[\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta) \right]^{-1} \right] = \left[\mathsf{I} - \mu(\mathsf{G} + \mathsf{G}^T) \right]^{-1} + \sigma^2 \, \mathsf{S}^{-1} \left(\sum_i R_i \, \mathsf{S}^{-1} \, R_i \right) \mathsf{S}^{-1}.$$

Plugging the above equality in (67) finally implies that

$$\begin{aligned} \Pi_{\mathsf{G}}^{\text{multi.}} &= \frac{a^2}{2} \mathbf{1}^T \, \mathbf{E}\left[\mathsf{I} - (\mathsf{M}_\theta \mathsf{G} + \mathsf{G}^T \mathsf{M}_\theta) \right]^{-1} \, \mathbf{1} \\ &= \frac{a^2}{2} \mathbf{1}^T \left[\mathsf{S}^{-1} + \sigma^2 \, \mathsf{S}^{-1} \left(\sum_i R_i \, \mathsf{S}^{-1} \, R_i \right) \mathsf{S}^{-1} \right] \mathbf{1}, \end{aligned} \tag{71}$$

recall that $\mathsf{S}^{-1} = \left[\mathsf{I} - \mu(\mathsf{G} + \mathsf{G}^T) \right]^{-1}$.

The proof is complete by comparing (66) and (71), that is

$$\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\sigma^2 \left\{ \left[K^T \sum_{i=1}^{n} R_i(\mathsf{S}^{-1} - \mathsf{I})R_i \ K \right] + \sum_{i=1}^{n} k_i^2 \parallel \deg_{in}(i) \parallel_2^2 \right\}$$
$$= \left(\frac{a^2}{2}\sigma^2\right) K^T \left\{ \sum_{i=1}^{n} R_i(\mathsf{S}^{-1} - \mathsf{I})R_i + \text{diag}[(\mathsf{G} \circ \mathsf{G})\mathbf{1}] \right\} K$$

since

$$K = \mathsf{S}^{-1}\mathbf{1} = [k_1 \ k_2 \ \cdots \ k_n]$$

and

$$\sum_{i=1}^{n} k_i^2 \parallel \deg_{in}(i) \parallel_2^2 = \mathbf{1}^T \mathsf{S}^{-1} \text{diag}\left(\sum_{j=1}^{n} g_{1j}^2, \sum_{j=1}^{n} g_{2j}^2, \cdots, \sum_{j=1}^{n} g_{nj}^2 \right) \mathsf{S}^{-1}\mathbf{1}$$
$$= K^T \text{diag}[(\mathsf{G} \circ \mathsf{G})\mathbf{1}] K.$$

□

Proof of Proposition 8. We prove each part separately as follows.

Part (i) The proof of the first part is immediate by (16). Note that k_i, for all i, increases by adding g_{ij}, assuming all the invertibility assumptions are preserved.

Part (ii) Let $0 < \alpha < 1$. Thus, the wight reduction is captured by $\alpha\mathsf{G}$, introducing a new network denoted by $\tilde{\mathsf{G}} = \alpha\mathsf{G}$. Using Lemma 2, since $\tilde{\mathsf{G}}^T\tilde{\mathsf{G}} = \alpha^2\mathsf{G}^T\mathsf{G}$ and $\tilde{R}_i = \alpha R_i$, for all i, thus we have:

$$\Pi_{\tilde{\mathsf{G}}}^{\text{multi.}} - \Pi_{\tilde{\mathsf{G}}}^{\text{bi.}} \Big|_{\tilde{\mathsf{G}}=\alpha\mathsf{G}} = \left(\frac{a^2}{2}\sigma^2\right) \alpha^2 \ \mathbf{1}^T \tilde{\mathsf{S}}^{-1} \left\{ \sum_{i=1}^{n} R_i \ \tilde{\mathsf{S}}^{-1} R_i \ - \mathsf{G}^T\mathsf{G} \right\} \tilde{\mathsf{S}}^{-1}\mathbf{1}$$
$$\leq \left(\frac{a^2}{2}\sigma^2\right) \alpha^2 \ \mathbf{1}^T \mathsf{S}^{-1} \left\{ \sum_{i=1}^{n} R_i \ \mathsf{S}^{-1} R_i \ - \mathsf{G}^T\mathsf{G} \right\} \mathsf{S}^{-1}\mathbf{1}$$
$$= \alpha^2 \left(\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} \right),$$

where the inequality follows because $\tilde{\mathsf{S}}^{-1} = (\mathsf{I} - \mu(\alpha\mathsf{G} + \alpha\mathsf{G}^T))^{-1}$ is increasing in α, thus $\tilde{\mathsf{S}}^{-1} \leq \mathsf{S}^{-1}$ (component-wise), completing the proof.

Part (iii) Since G is symmetric (i.e. $\mathsf{G} = \mathsf{G}^T$) it has n distinct eigenvalues and is thus diagonalizable. Therefore, there exists an invertible matrix V so that $V^{-1}\mathsf{G}V = \Lambda$, where Λ is a diagonal matrix with the eigenvalues of G on its diagonal, and columns of V are the eigenvectors of G. Therefore, $\mathsf{G} = V\Lambda V^{-1}$. In addition, the eigenvectors can be chosen to form an orthonormal basis of $\mathbb{R}^n$, meaning that V is set to be orthonormal. That is $V^T V = V V^T = V^{-1} V = \mathsf{I}$. Moreover,

$$\mathsf{S}^{-1} = \sum_{i=0}^{\infty}(2\mu\ \mathsf{G})^i = \sum_{i=0}^{\infty}\left(2\mu\ V\Lambda V^{-1}\right)^i = V\left(\sum_{i=0}^{\infty}(2\mu\ \Lambda)^i\right)V^{-1} = V\Lambda_1 V^{-1}, \tag{72}$$

where Λ_1 is a diagonal matrix where its k−th element is $\frac{1}{1-2\mu\lambda_k}$, (note that $2\mu\lambda_k \leq 2\mu\lambda_{max} < 1$, for all k).

By following similar argument as in (72), one can also show:

$$\begin{aligned} \mathsf{S}^{-1}\ \mathsf{G}^T\mathsf{G}\ \mathsf{S}^{-1} &= V\Lambda_1^2\Lambda^2 V^{-1} \\ \mathsf{S}^{-1}\left(\sum_{i=1}^{n} R_i\right)\mathsf{S}^{-1}\left(\sum_{i=1}^{n} R_i\right)\mathsf{S}^{-1} &= V\Lambda_1^3(2\Lambda)^2 V^{-1}. \end{aligned} \tag{73}$$

Note that $\sum_{i=1}^{n} R_i = \mathsf{G} + \mathsf{G}^T = 2\mathsf{G}$, (since $\mathsf{G} = \mathsf{G}^T$).

Thus, (73) yields:

$$\Xi \triangleq \mathsf{S}^{-1}\,(2\mathsf{G})\ \mathsf{S}^{-1}\,(2\mathsf{G})\mathsf{S}^{-1} - \ \mathsf{S}^{-1}\ \mathsf{G}^T\mathsf{G}\ \mathsf{S}^{-1} = V\left(\underbrace{\Lambda_1^3(2\Lambda)^2 - (\Lambda_1\Lambda)^2}_{\text{diagonal}}\right)V^{-1}. \tag{74}$$

The above equality characterizes eigenvalues of Ξ in terms of $\lambda_1, \lambda_2, \cdots, \lambda_n$. That is, k−th eigenvalue of Ξ is equal to $f(\lambda_k) = \lambda_k^2\ \frac{3+2\mu\lambda_k}{(1-2\mu\lambda_k)^3}$. Further, it can be easily shown $f(\lambda)$ is increasing and convex in λ.

Next, using Lemma 2, we have

$$\begin{aligned}
\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} &= \left(\frac{a^2}{2}\sigma^2\right) \ \mathbf{1}^T \mathsf{S}^{-1} \left\{ \sum_{i=1}^{n} R_i \ \mathsf{S}^{-1} R_i \ - \mathsf{G}^T \mathsf{G} \right\} \mathsf{S}^{-1} \mathbf{1} \\
&\leq \left(\frac{a^2}{2}\sigma^2\right) \ \mathbf{1}^T \mathsf{S}^{-1} \left\{ \left(\sum_{i=1}^{n} R_i \right) \mathsf{S}^{-1} \left(\sum_{i=1}^{n} R_i \right) - \mathsf{G}^T \mathsf{G} \right\} \mathsf{S}^{-1} \mathbf{1} \\
&= \left(\frac{a^2}{2}\sigma^2\right) \mathbf{1}^T \Xi \mathbf{1}.
\end{aligned}$$

Finally, the proof of the proposition is immediate by employing Lemma 6 and the fact that $\mathbf{1}^T\mathbf{1} = n$.

Part (iv) To prove this part, we only use the second term in (16). Let d_i denote agent i's in degree in G (thus, by assumption, $\sum_j g_{ij} = g d_i$). Suppose agent i_{max} has the maximum in degree in G. Thus: (i) by definition, $k_{i_{max}} > 1 + \mu(d_{i_{max}} g)$. (ii) $\| \text{deg}_{in}(i_{max}) \|_2^2 = g^2 d_{i_{max}}$. Hence, (i) together with (ii) give:

$$\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} \geq \left(\frac{a^2}{2}\sigma^2\right) \sum_{i=1}^{n} k_i^2 \, \| \text{deg}_{in}(i) \|_2^2 \geq \left(\frac{a^2}{2}\sigma^2\right) (\mu \, d_{i_{max}} g)^2 g^2 d_{i_{max}} = O(d_{i_{max}}^3 g^4),$$

completing the proof.

□

Proof of Proposition 12. Using Lemma 2, we have

$$\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\sigma^2 \left\{ \underbrace{K^T \sum_{i=1}^{n} R_i (\mathsf{S}^{-1} - \mathsf{I}) R_i \ K}_{\triangleq T_1} + \underbrace{\sum_{i=1}^{n} k_i^2 \, \| \text{deg}_{in}(i) \|_2^2}_{\triangleq T_2} \right\}. \tag{75}$$

We analyze T_1 and T_2, separately, as follows. Let G denote the adjacency matrix of a star network with n nodes, where agent 1 located at the center of it and the rest are at the periphery.

Given the definitions of R_i, for all i (that is, $R_i \triangleq E_i \mathsf{G} + \mathsf{G}^T E_i$, where E_i is the matrix with only ith diagonal set as 1 and other entries as zero) for any non-negative matrix A

we clearly have:

$$\sum_{i=1}^{n} R_i \; A \; R_i \leq \left(\sum_{i=1}^{n} R_i \right) \; A \; \left(\sum_{i=1}^{n} R_i \right),$$

where here ($\leq$) is for component-wise comparison. Since S^{-1} is invertible and well defined, thus $\mathsf{S}^{-1} - \mathsf{I}$ is component-wise non-negative. Thus, using the above inequality we have:

$$\sum_{i=1}^{n} R_i \; (\mathsf{S}^{-1} - \mathsf{I}) \; R_i \leq \left(\sum_{i=1}^{n} R_i \right) \; (\mathsf{S}^{-1} - \mathsf{I}) \; \left(\sum_{i=1}^{n} R_i \right). \tag{76}$$

Focusing on directed star networks (with no parallel links) where node 1 is located at the center, the right-hand-side of the above inequality is achievable for star-inward networks in which

$$\sum_{i=1}^{n} R_i = R_1 \triangleq \kappa \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}.$$

Recall that we have assumed $g_{ij} \in \{0, \kappa\}$ for all i, j, where $\kappa > 0$.

Moreover, since by definition $K \triangleq \mathsf{S}^{-1}\mathbf{1} = [k_1 \; k_2 \; \cdots \; k_n]^T$ is the same for all of these star networks, thus, using (76), we also have

$$K^T \left(\sum_{i=1}^{n} R_i \; (\mathsf{S}^{-1} - \mathsf{I}) \; R_i \right) K \leq K^T \left(\sum_{i=1}^{n} R_i \right) \; (\mathsf{S}^{-1} - \mathsf{I}) \; \left(\sum_{i=1}^{n} R_i \right) K, \tag{77}$$

that means the right-hand-side is achievable with the star-inward network. Thus, in (75), maximum of T_1 is obtained star-inward networks.

We next complete the proof by considering T_2 in (75). Again, note that in all of these star networks $K = \mathsf{S}^{-1}\mathbf{1} = [k_1 \; k_2 \; \cdots \; k_n]^T$ is the same, ($k_1 > k_2 = k_3 = \cdots = k_n$). But in star inward network

$$T_2 = \kappa^2 (n-1) k_1^2. \tag{78}$$

We claim T_2 is maximum in the case of a star-inward network. To prove this, consider another star network in which agent 1 obtains externality from n_1 neighbors, and n_2

periphery nodes obtain externality from agent 1, thus $n_1 + n_2 = n-1$. In this star network $T_2 = \kappa^2 n_1 k_1^2 + \kappa^2 n_2 k_2^2 = \kappa^2(n_1 k_1^2 + n_2 k_2^2)$, that is lower than (78), because $k_1 > k_2$.

$\square$

Proof of Proposition 13. To prove the proposition, we first show that if $\alpha_n = \frac{1}{\sqrt{(n-1)(\mu^2+\sigma^2+\epsilon)}}$, where ϵ is positive and constant, then $\left(\mathsf{I} - \mu\left(\mathsf{G}(n) + \mathsf{G}(n)^T\right) + \sigma^2 \mathsf{G}(n)^T \mathsf{G}(n)\right)^{-1}$ is well-defined and non-negative. For ease of illustration denote $G(n) \triangleq \mu\left(\mathsf{G}(n) + \mathsf{G}(n)^T\right) - \sigma^2 \mathsf{G}(n)^T \mathsf{G}(n)$. To prove the invertibility, it is enough to show that $\rho(G(n)) < 1$, i.e., the maximum eigenvalue of $G(n)$, in absolute value, is less than 1. Note that $G(n)$ is symmetric, thus all of its eigenvalues are real. In the sequel, fix n. It is clear that:

$$\mathsf{G}(n) + \mathsf{G}(n)^T = \alpha_n \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix},$$

and

$$\mathsf{G}(n)^T \mathsf{G}(n) = \alpha_n^2 \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 1 \end{bmatrix}.$$

Let λ be an eigenvalue of $G(n)$. By definition, $G(n)x = \lambda x$ does have a solution in x. Due to the symmetry, x has the following form, $x^T = [a \; b \; b \; \cdots \; b]$. Thus, due to the definition of $G(n)$ and the above two (matrix-from) equalities, $G(n)x = \lambda x$ yields

$$\begin{aligned} \lambda a &= \alpha_n \mu (n-1) b, \\ \lambda b &= \alpha_n \mu a + \sigma^2 \alpha_n^2 (n-1) b. \end{aligned}$$

Thus, by a little algebra, we have

$$\lambda^2 - \underbrace{\alpha_n^2 \sigma^2 (n-1)}_{\triangleq B} \lambda - \underbrace{\alpha_n^2 \mu^2 (n-1)}_{\triangleq C} = 0. \tag{79}$$

that is, $\lambda^2 - B\lambda - C = 0$. Now, note that $0 < C < B$ and $B + C < 1$, where the later is ensured by $\alpha_n = \frac{1}{\sqrt{(n-1)(\mu^2+\sigma^2+\epsilon)}}$, where $\epsilon > 0$ and is constant. Therefore, the roots of (79) are (in absolute value) less than 1, that is desired. This also implies $\left(\mathsf{I} - \mu\left(\mathsf{G}(n) + \mathsf{G}(n)^T\right)\right)^{-1}$ is well-defined and non-negative. Next, we continue to complete the proof. By definition, we have:

$$\begin{aligned} k_1(n) = \left[\left(\mathsf{I} - \mu\left(\mathsf{G}(n) + \mathsf{G}(n)^T\right)\right)^{-1} \mathbf{1}\right]_1 &> \left[\left(\mathsf{I} + \mu\left(\mathsf{G}(n) + \mathsf{G}(n)^T\right)\right)\mathbf{1}\right]_1 \\ &= 1 + \mu\alpha_n(n-1). \end{aligned}$$

From the above inequality we obtain:

$$k_1(n) = O\left(\sqrt{n}\right).$$

Finally, using Lemma 2, for all n, we have:

$$\begin{aligned} \Pi^{\text{multi.}}_{\mathsf{G}(n)} - \Pi^{\text{bi.}}_{\mathsf{G}(n)} &> \frac{a^2}{2}\sigma^2 \left\{ \sum_{i=1}^{n} k_i^2 \parallel \deg_{in}(i) \parallel_2^2 \right\} \\ &= \frac{a^2}{2}\sigma^2 \; k_1^2(n) \; \alpha_n^2 \; (n-1) \\ &= \frac{a^2}{2}\left(\frac{\sigma^2}{\mu^2 + \sigma^2 + \epsilon}\right) k_1^2(n). \end{aligned}$$

The last equality implies that as n grows, $\Pi^{\text{multi.}}_{\mathsf{G}(n)} - \Pi^{\text{bi.}}_{\mathsf{G}(n)}$ tends to infinity, as $k_1(n)$ goes to infinity.

□

Proof of Proposition 14. Using Proposition 12, proving this result is equivalent to show the following.

Equivalent Proposition: Consider the sequence of Star-inward networks $\{\mathsf{G}(n)\}_{n=2}^{\infty}$, where $\beta_n \triangleq g_{12} = g_{13} = \cdots = g_{1n} = \frac{1}{(n-1)\sqrt{\mu^2+\sigma^2+\epsilon}}$, where $\epsilon > 0$ is a constant, and the rest of entries are zero. Then we have:

$$\Pi^{\text{multi.}}_{\mathsf{G}(n)} - \Pi^{\text{bi.}}_{\mathsf{G}(n)} \longrightarrow 0.$$

We first note since $\beta_n = \frac{1}{(n-1)\sqrt{\mu^2+\sigma^2+\epsilon}}$, as shown in Proposition 13,

$\left(\mathsf{I}-\mu\left(\mathsf{G}(n)+\mathsf{G}(n)^T\right)-\sigma^2\mathsf{G}(n)^T\mathsf{G}(n)\right)^{-1}$ is well-defined and non-negative, for all n.

Now, fix $n>1$, and denote

$$R_1 \triangleq \beta_n \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} = \mathsf{G}+\mathsf{G}^T,$$

thus, for any $t>0$,

$$R_1^{2t+1} = \beta_n^{2t+1} \begin{bmatrix} 0 & (n-1)^t & \cdots & (n-1)^t \\ (n-1)^t & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (n-1)^t & 0 & \cdots & 0 \end{bmatrix}, \tag{80}$$

and

$$R_1^{2t} = \beta_n^{2t} \begin{bmatrix} (n-1)^t & 0 & \cdots & 0 \\ 0 & (n-1)^{t-1} & \cdots & (n-1)^{t-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & (n-1)^{t-1} & \cdots & (n-1)^{t-1} \end{bmatrix}. \tag{81}$$

We first compute Bonacich centrality measures of agents in $\mathsf{G}+\mathsf{G}^T$. By definition, Bonacich centrality vector is given by: $[k_1\ k_2\ \cdots\ k_2]^T = \left(\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T)\right)^{-1}\mathbf{1}$. Due to the symmetry and the invertibility, $\left(\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T)\right)^{-1} = \sum_{t\geq 0}\left(\mu(\mathsf{G}+\mathsf{G}^T)\right)^t = \sum_{t\geq 0}(\mu R_1)^t$. Thus, by simple algebra, it follows that

$$k_1 = 1+\beta_n\mu(n-1)+\sum_{t=2}^{\infty}(\beta_n\mu)^t\,(n-1)^{\lceil \frac{t}{2}\rceil} = O(1) = \text{constant}, \tag{82}$$

$$k_2 = 1+\beta_n\mu+\sum_{t=2}^{\infty}(\beta_n\mu)^t\,(n-1)^{\lfloor \frac{t}{2}\rfloor} = O(1) = \text{constant}, \tag{83}$$

where the last equality in the above two equalities are followed since n is large and $\beta_n =$

$\frac{1}{(n-1)\sqrt{\mu^2+\sigma^2+\epsilon}}$. Now, using Proposition 2, we have

$$\Pi_{\mathsf{G}}^{\text{multi.}} - \Pi_{\mathsf{G}}^{\text{bi.}} = \frac{a^2}{2}\sigma_\delta^2 \left\{ \underbrace{K^T \sum_{i=1}^{n} R_i(\mathsf{S}^{-1} - \mathsf{I})R_i \ K}_{\triangleq T_1} \ + \ \underbrace{\sum_{i=1}^{n} k_i^2 \parallel \deg_{in}(i) \parallel_2^2}_{\triangleq T_2} \right\}, \tag{84}$$

thus, in this set up, $T_2 = k_1^2 \beta_n^2 (n-1) \to 0$. This is because $\beta_n = \frac{1}{(n-1)\sqrt{\mu^2+\sigma^2+\epsilon}}$ and $k_1 = O(1)$. In the sequel of the proof, according to (84), we focus on T_1. According to the definitions of R_i (see proof of Proposition 2), we have:

$$\begin{aligned}
T_1 &= K^T \sum_{i=1}^{n} R_i(\mathsf{S}^{-1} - \mathsf{I})R_i \ K, \\
&= K^T R_1(\mathsf{S}^{-1} - \mathsf{I})R_1 \ K, \\
&= K^T R_1 \left(\sum_{t=1}^{\infty} (\mu R_1)^t \right) R_1 \ K, \\
&= K^T \left(\sum_{t=1}^{\infty} \mu^t \ R_1^{t+2} \right) K \\
&= \underbrace{K^T \left(\sum_{t=1}^{\infty} \mu^{2t-1} \ R_1^{2t+1} \right) K}_{\triangleq T_{odd}} + \underbrace{K^T \left(\sum_{t=2}^{\infty} \mu^{2t-2} \ R_1^{2t} \right) K}_{\triangleq T_{even}}.
\end{aligned} \tag{85}$$

Recall that $K^T = [k_1 \ k_2 \ \cdots \ k_2]$. Further, using (80) and (81), for any t, we have:

$$\begin{aligned}
K^T R_1^{2t+1} K &= \beta_n^{2t+1} \ (2k_1 k_2 (n-1)^{t+1}), \\
K^T R_1^{2t} K &= \beta_n^{2t} \ (k_1^2 + (n-1)k_2^2) \ (n-1)^t.
\end{aligned}$$

Plugging the above equalities into (85), and approaching n to infinity, we obtain:

$$T_{odd} = 2k_1k_2\sum_{t=1}^{\infty}\mu^{2t-1}(n-1)^{t+1}\beta_n^{2t+1} \to 0,$$
$$T_{even} = (k_1^2+(n-1)k_2^2)\sum_{t=2}^{\infty}\mu^{2t-2}(n-1)^{t}\beta_n^{2t} \to 0,$$

That is because $O(k_1) = O(k_2) = 1$ and $\beta_n = O\left(\frac{1}{n}\right)$. Thus $T_1 \to 0$, that completes the proof.

□

Proof of Proposition 9. To prove the result we show $\Pi^{\text{multi.}}_{\mathsf{G}(n)} = \Pi^{\text{bi.}}_{\mathsf{G}(n)} = O(n)$. Using (77) we have:

$$\begin{aligned}
\mathbf{1}^T\mathsf{S}^{-1}\sum_i R_i\mathsf{S}^{-1}R_i\mathsf{S}^{-1}\mathbf{1} &= \zeta^2\mathbf{1}^T\left(\sum_i R_i\mathsf{S}^{-1}R_i\right)\mathbf{1}\\
&\le \zeta^2\mathbf{1}^T\left(\sum_i R_i\right)\mathsf{S}^{-1}\left(\sum_i R_i\right)\mathbf{1}\\
&= \zeta^2\mathbf{1}^T\left(\mathsf{G}+\mathsf{G}^T\right)\mathsf{S}^{-1}\left(\mathsf{G}+\mathsf{G}^T\right)\mathbf{1}\\
&= \zeta^2\tau^2\mathbf{1}^T\mathsf{S}^{-1}\mathbf{1}\\
&= \zeta^3\tau^2 n\\
&= O(n).
\end{aligned}$$

Thus, the above inequality along with (71) imply that

$$\begin{aligned}
\Pi^{\text{bi.}}_{\mathsf{G}(n)} \le \Pi^{\text{multi.}}_{\mathsf{G}(n)} &= \frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1}+\sigma^2\,\mathsf{S}^{-1}\left(\sum_i R_i\,\mathsf{S}^{-1}\,R_i\right)\mathsf{S}^{-1}\right]\mathbf{1}\\
&\le \frac{a^2}{2}\left[\zeta n+\sigma^2\zeta^3\tau^2 n\right]\\
&= O(n).
\end{aligned}$$

In addition, using Proposition 7, since the in degree at each node is the same, it clears that $\Pi^{\text{bi.}}_{\mathsf{G}(n)} = O(n)$. Therefore, $\lim_{n\to\infty}\frac{\Pi^{\text{multi.}}_{\mathsf{G}(n)}}{\Pi^{\text{bi.}}_{\mathsf{G}(n)}} = O(1)$.

To complete the proof, we need to guaranty the invertibility of $T = \left(\mathsf{I}-\mu(\mathsf{G}+\mathsf{G}^T)-\sigma^2\mathsf{G}^T\mathsf{G}\right)$. Choosing small enough τ (independent of n) will ensure this, if we show $\mathsf{G}^T\mathsf{G}$ dose not blow up as n grows to infinity. Assuming $\mathsf{G}+\mathsf{G}^T$ is k-regular (k is finite), we obtain

$$\sum_j [\mathsf{G}^T\mathsf{G}]_{ij} = \text{number of walks of length 2 starting from } i = k^2,$$

which is independent of n. Meaning that by choosing small τ (independent of n) T can be diagonally dominant and of course invertible. □

The following result shows the above observation is also true in **cycle (clock-wise) networks** that may not necessarily have a balanced structure.

Proposition 16. *Consider the sequence of (clock-wise) ring/cycle networks* $\{\mathsf{G}(n)\}_{n=2}^{\infty}$*, where weights are either zero or one. Let* $2\mu+\sigma^2<1$ *and* $\mu<\frac{1}{4}$*. As n grows:*

$$\lim_{n\to\infty}\frac{\Pi^{multi.}_{\mathsf{G}(n)}}{\Pi^{bi.}_{\mathsf{G}(n)}} = O(1).$$

Proof of Proposition 16. Fix n. We first note that since

$$\mathsf{G}(n) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

thus,

$$\mathsf{G}(n)^T\mathsf{G}(n) = \mathsf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

and

$$\mathsf{G}(n)+\mathsf{G}(n)^T=\begin{bmatrix}0 & 1 & 0 & 0 & \cdots & 1\\ 1 & 0 & 1 & 0 & \cdots & 0\\ 0 & 1 & 0 & 1 & \cdots & 0\\ \vdots & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 & 1\\ 1 & 0 & \cdots & 0 & 1 & 0\end{bmatrix}.$$

Therefore, when $2\mu+\sigma^2<1$, then $\left(\mathsf{I}-\mu\left(\mathsf{G}(n)+\mathsf{G}(n)^T\right)-\sigma^2\mathsf{G}(n)^T\mathsf{G}(n)\right)^{-1}$ is well-defined and non-negative[64]. Furthermore, Bonacich centrality measures of agents in $\mathsf{G}(n)+\mathsf{G}(n)^T$ (by definition) is given by: $[k_1\ k_2\ \cdots\ k_2]^T=\mathsf{S}(n)^{-1}\mathbf{1}=\left(\mathsf{I}-\mu(\mathsf{G}(n)+\mathsf{G}(n)^T)\right)^{-1}\mathbf{1}$. Due to the symmetry and the invertibility we have $k_1=\cdots=k_n=k$, and since $\mu<\frac{1}{4}$, by simple algebra, we have $1\le k<2$. That is, $k=O(1)$.

As shown in Proposition 2

$$\Pi^{\text{bi.}}_{\mathsf{G}(n)}=\frac{a^2}{2}\mathbf{1}^T\left[\mathsf{S}^{-1}+\sigma^2\,\mathsf{S}(n)^{-1}\underbrace{\mathsf{G}(n)^T\mathsf{G}(n)}_{=\mathsf{I}}\mathsf{S}(n)^{-1}\right]\mathbf{1}=\frac{a^2}{2}\left[\mathbf{1}^T\mathsf{S}^{-1}\mathbf{1}^T+\sigma^2\,\mathbf{1}^T\mathsf{S}(n)^{-1}\ \mathsf{S}(n)^{-1}\mathbf{1}\right]$$
$$=\frac{a^2}{2}\left[kn+\sigma^2k^2n\right]=O(n). \tag{86}$$

Next, we show $\Pi^{\text{multi.}}_{\mathsf{G}(n)}=O(n)$.

As shown in Lemma 2 (see Eq. (71)) we have:

$$\Pi^{\text{multi.}}_{\mathsf{G}(n)}=\frac{a^2}{2}\mathbf{1}^T\ \mathbf{E}\left[\mathsf{I}-(\mathsf{M}_\delta\mathsf{G}(n)+\mathsf{G}(n)^T\mathsf{M}_\delta)\right]^{-1}\ \mathbf{1}$$
$$=\frac{a^2}{2}\left[\underbrace{\mathbf{1}^T\mathsf{S}(n)^{-1}\mathbf{1}}_{T_1}+\underbrace{\sigma^2\ \mathbf{1}^T\ \mathsf{S}(n)^{-1}\left(\sum_i R_i\ \mathsf{S}(n)^{-1}\ R_i\right)\mathsf{S}(n)^{-1}\mathbf{1}}_{T_2}\right],$$

First we note that, by definition, $T_1=kn=O(n)$. Next we show $T_2=O(n)$. Notice that

[64] This is because $\mathsf{I}-\mu\left(\mathsf{G}(n)+\mathsf{G}(n)^T\right)-\sigma^2\mathsf{G}(n)^T\mathsf{G}(n)$ becomes diagonally dominant (when $2\mu+\sigma^2<1$).

$\mathsf{S}(n)^{-1}\mathbf{1} = k\mathbf{1}$ and $\mathbf{1}^T\mathsf{S}(n)^{-1} = k\mathbf{1}^T$. Thus,

$$T_2 = \sigma^2 k^2 \ \mathbf{1}^T \left(\sum_i R_i \ \mathsf{S}(n)^{-1} \ R_i \right) \mathbf{1}.$$

Therefore,

$$T_2 \leq \sigma^2 k^2 \ \ \mathbf{1}^T \left[\left(\sum_i R_i \right) \mathsf{S}(n)^{-1} \left(\sum_i R_i \right) \right] \mathbf{1} = 2\sigma^2 k^2 \ \ \mathbf{1}^T \mathsf{S}(n)^{-1} \mathbf{1} = 2\sigma^2 k^3 n. \tag{87}$$

Also since $\mathsf{S}(n)^{-1} \geq \mathsf{I}$, thus

$$T_2 \geq \sigma^2 k^2 \ \mathbf{1}^T \left(\sum_i R_i \ R_i \right) \mathbf{1} \geq \sigma^2 k^2 \ \ \mathbf{1}^T \mathbf{1} = \sigma^2 k^2 n.$$

Eq. (87) along with (88) imply that $T_2 = O(n)$. Thus, $\Pi^{\text{multi.}}_{\mathsf{G}(n)} = O(n)$. This along with (86) complete the proof.

□

Proof of Proposition 10. Without loss of generality, for the ratio analysis, one can set $\frac{a^2}{2} = 1$. Thus:

$$\begin{aligned}
\Pi^{multi.}_{\mathsf{G}} &= \mathbf{1}^T K + \sigma^2 \ K^T \left(\sum_i R_i \ \mathsf{S}^{-1} \ R_i \right) K \\
&\leq \mathbf{1}^T K + \sigma^2 \ K^T \left(\sum_i R_i \right) \ \mathsf{S}^{-1} \ \left(\sum_i R_i \right) K \\
&= \mathbf{1}^T K + \sigma^2 \ K^T \left(\mathsf{G} + \mathsf{G}^T \right) \ \mathsf{S}^{-1} \ \left(\mathsf{G} + \mathsf{G}^T \right) K \\
&= \mathbf{1}^T K + \sigma^2 \ \mathbf{1}^T \left(\mathsf{G} + \mathsf{G}^T \right)^2 \ \mathsf{S}^{-3} \mathbf{1}.
\end{aligned}$$

Let λ_{max} be the maximum eigenvalue of $\mathsf{G} + \mathsf{G}^T$. Note that invertibility ensures $\mu\lambda_{max} < 1$. Since $\mathsf{G} + \mathsf{G}^T$ is symmetric, it is diagonalizable. There exists orthonormal matrices P and P^T such that $PP^{-1} = PP = PP^T = \mathsf{I}$, and $P(\mathsf{G} + \mathsf{G}^T)P^T = \text{diag}(\lambda_1, \cdots, \lambda_n)$, where λ_i are eigenvalues of $\mathsf{G} + \mathsf{G}^T$.

Using the digitalization of $\mathsf{G}+\mathsf{G}^T$ yields

$$(\mathsf{G}+\mathsf{G}^T)^2\mathsf{S}^{-3} = P^T \text{diag}(f(\lambda_1),\cdots,f(\lambda_n))P$$

where $f(\lambda) \triangleq \frac{\lambda^2}{(1-\mu\lambda)^3}$, which is increasing and convex in λ.

Since $f(\lambda)$ is increasing in λ, thus the maximum eigenvalue of $(\mathsf{G}+\mathsf{G}^T)^2\mathsf{S}^{-3}$ becomes $f(\lambda_{max})$. This implies that $\mathbf{1}^T\left(\mathsf{G}+\mathsf{G}^T\right)^2\ \mathsf{S}^{-3}\mathbf{1} \le nf(\lambda_{max})$. Therefore, we obtain:

$$\Pi_{\mathsf{G}}^{multi.} \le \|K\|_1 + \sigma^2\ n\ f(\lambda_{max}). \tag{88}$$

To wrap up the proof, recall that

$$\Pi_{\mathsf{G}}^{bi.} = \|K\|_1 + \sigma^2 K^T\mathsf{G}^T\mathsf{G}K \ge \|K\|_1. \tag{89}$$

Since $\|K\|_1 \ge n$, the proof is complete. □

Proof of Proposition 11. Again, without loss of generality, for the ratio analysis, one can assume $\frac{a^2}{2} = 1$. Thus:

$$\begin{aligned}\Pi_{\mathsf{G}}^{multi.} &= \mathbf{1}^T K + \sigma^2\ K^T\left(\sum_i R_i\ \mathsf{S}^{-1}\ R_i\right)K\\ &\ge \mathbf{1}^T K + \sigma^2\ K^T\left(\sum_i R_i^2\right)K\\ &= \mathbf{1}^T K + \sigma^2\ K^T\left(\mathsf{G}^T\mathsf{G}+\tilde{M}\right)K,\end{aligned}$$

where $\tilde{M} = \text{diag}(\|d_1\|_2^2, \|d_2\|_2^2,\cdots,\|d_n\|_2^2)$. Therefore,

$$\begin{aligned}\frac{\Pi_{\mathsf{G}}^{multi.}}{\Pi_{\mathsf{G}}^{bi.}} &\ge \frac{\mathbf{1}^T K + \sigma^2\ K^T\left(\mathsf{G}^T\mathsf{G}+\tilde{M}\right)K}{\mathbf{1}^T K + \sigma^2\ K^T\mathsf{G}^T\mathsf{G}K}\\ &= 1 + \frac{\sigma^2\ K^T\tilde{M}K}{\mathbf{1}^T K + \sigma^2\ K^T\mathsf{G}^T\mathsf{G}K}\\ &\ge 1 + \frac{\sigma^2\|\text{diag}(K)\mathsf{G}\|_F^2}{\|K\|_1 + \|K\|_2^2}\end{aligned}$$

where the last inequality follows by noting that the maximum eigenvalue of $\sigma^2\mathsf{G}^T\mathsf{G}$ is less than 1, and therefore $\sigma^2 K^T \mathsf{G}^T \mathsf{G} K \leq K^T K = \|K\|_2^2$.

□

F Appendix: Different centrality orders in $\mathsf{G}+\mathsf{G}^T$ and $\mathcal{G}+\mathcal{G}^T$

For ease of notation we first have the following definition.

Definition 6 (Modified Network: $\mathcal{G}$). *Given the adjacency matrix* G *and the mean and variance of the (virtual) types, respectively* μ_ψ *and* σ_ψ^2, *the* modified network $\mathcal{G}$ *is defined as*

$$\mathcal{G} \triangleq \mu_\psi \mathsf{G} + \frac{\sigma_\psi^2}{2b} \mathsf{G}^T \mathsf{G}. \tag{90}$$

The modified network for the first-best allocation is similarly defined via using μ *and* σ^2 *instead of* μ_ψ *and* σ_ψ^2*, respectively.*

A quick comparison between Propositions 1 and 3 shows that multilateral and bilateral contracts may *not* necessarily induce the same order of allocations. To see this, Let $b = 1$ and $\theta_i = \bar{\theta}$, for all i. Then, Proposition 1 implies $\mathbf{x}^{SB}_{\text{multilateral}} = a[\mathsf{I} - \bar{\theta}(\mathsf{G}+\mathsf{G}^T)]^{-1}\mathbf{1}$, that is proportional to the Bonacich centrality measure in $\mathsf{G}+\mathsf{G}^T$. However, Proposition 3 implies $\mathbf{x}^{SB}_{\text{bilateral}} = a(\Delta\theta\ \mathsf{G}+\mathsf{I})[\mathsf{I}-(\mathcal{G}+\mathcal{G}^T)]^{-1}\mathbf{1}$, where $\mathcal{G} \triangleq \mu_\psi \mathsf{G} + \frac{\sigma_\psi^2}{2b}\mathsf{G}^T\mathsf{G}$. That is proportional to the Bonacich centrality measure in $\mathcal{G}+\mathcal{G}$.

In this section we provide a condition under which centrality orders in the modified and the original networks are always diffident.

In general, the form of the modified network (see Def. 6) suggests that when $\mathsf{G}^T\mathsf{G}$ is sufficiently different from G, even with a small amount of uncertainty (i.e. σ^2), a change on the central agents in $\mathsf{G}+\mathsf{G}^T$ in compare to $\mathcal{G}+\mathcal{G}^T$ may happen. In other words, in networks with high second order of connectivity, captured by $\mathsf{G}^T\mathsf{G}$, a small amount of uncertainty may make a change on the centrality orders of agents in the original network.

Following the above intuition, in the next proposition, we present a sufficient condition under which centrality orders in $\mathsf{G}+\mathsf{G}^T$ and $\mathcal{G}+\mathcal{G}^T$ are different. For ease of exposition, let us define $d_{\mathsf{G}}^{out}(i)$ denoting the out-degree of agent i in G, i.e., $d_{\mathsf{G}}^{out}(i) = |\{k : g_{ki} > 0, k = 1, 2, \cdots, n\}|$.

Proposition 17. *Let agent* i *and* j *be both central in* $\mathsf{G}+\mathsf{G}^T$*. If* $d_{\mathsf{G}}^{out}(i) > d_{\mathsf{G}}^{out}(j) = 0$*, then Bonacich centrality of agent* i *is strictly higher than agent* j *in* $\mathcal{G}+\mathcal{G}^T$*.*

Proof of Proposition 17. By definition, we have

$$\frac{1}{b}(\mathsf{G}^T\mathsf{G})_{rs} = \frac{1}{b}\sum_{k=1}^{n} \tilde{g}_{rk}\; g_{ks} = \frac{1}{b}\sum_{k=1}^{n} g_{kr}\; g_{ks}, \tag{91}$$

From (91) we obtain for any r, s:

$$\begin{aligned}
&(\mathsf{G}^T\mathsf{G})_{rr} = \frac{1}{2}\sum_{k=1}^{n} g_{kr}^2 > 0, &&\text{if } d_{\mathsf{G}}^{out}(r) > 0.\\
&(\mathsf{G}^T\mathsf{G})_{rs} = (\mathsf{G}^T\mathsf{G})_{sr} = 0, &&\text{if } d_{\mathsf{G}}^{out}(r) \times d_{\mathsf{G}}^{out}(s) = 0.\\
&(\mathsf{G}^T\mathsf{G})_{rs} = (\mathsf{G}^T\mathsf{G})_{sr} > 0, &&\text{if } \exists\ k \text{ s.t. } g_{kr} > 0 \text{ and } g_{ks} > 0.
\end{aligned} \tag{92}$$

Bonacich centrality measure of an agent in a given network counts the total number of (suitably weighted) walks of different length starting from the agent in the network. Next, consider agent i and j such that $d_{\mathsf{G}}^{out}(i) > d_{\mathsf{G}}^{out}(j) = 0$. In $\mathsf{G}^T\mathsf{G}$: since $d_{\mathsf{G}}^{out}(j) = 0$, thus there is no walk starting from agent j, whereas, because $d_{\mathsf{G}}^{out}(i) > 0$, there exits walks (at least self-loops) starting from agent i.

Therefore, because: (i) $d_{\mathsf{G}}^{out}(i) > d_{\mathsf{G}}^{out}(j) = 0$, (ii) agent i and j are both equivalently central in $\mathsf{G}+\mathsf{G}^T$, and (iii) $\mathcal{G}+\mathcal{G}^T = \mu(\mathsf{G}+\mathsf{G}^T) + \frac{\sigma^2}{b}\mathsf{G}^T\mathsf{G}$, it follows that Bonacich centrality of agent i is strictly higher than agent j in $\mathcal{G}+\mathcal{G}^T$. □

The above proposition intuitively states that having a difference in the out-degrees of two central agents in $\mathsf{G}+\mathsf{G}^T$, might be crucial to make a change in their centralities with respect to the modified network. This proposition is substantiated by the following example.

Example 4 (In-ward-out-ward stars)**.** *Consider the following network* G *with* $2d+2$ *agents*[65] *wherein* $d_{\mathsf{G}}^{out}(1) = d+1$, $d_{\mathsf{G}}^{out}(i) = 1$, $d+3 \le i \le 2d+2$, *and* $d_{\mathsf{G}}^{out}(2) = d_{\mathsf{G}}^{out}(j) = 0$, $3 \le j \le d+2$, *as it is depicted in Fig. 9. In the modified network (depicted in Fig. 9) the impact of* $\mathsf{G}^T\mathsf{G}$ *is drawn by red lines. Since* $d_{\mathsf{G}}^{out}(1) = d+1$ *and* $d_{\mathsf{G}}^{out}(i) = 1$, $d+3 \le i \le 2d+2$, *thus due to* $\mathsf{G}^T\mathsf{G}$ *node* 1 *and node* i, $d+3 \le i \le 2d+2$, *have, respectively,* $d+1$ *and* 1 *self-loops. Moreover, node* 1 *and* i, $d+3 \le i \le 2d+2$, *both feed agent* 2, *thus agents* 1 *and* i *are connected in* $\mathcal{G}$ *due to* $\mathsf{G}^T\mathsf{G}$*. Now, applying Theorem 17, it follows that agents* 1 *and* 2 *are both, equivalently, central*

[65]Suppose $g_{ij} \in \{0, k\}$, for any i, j, where $k > 0$ is sufficiently small enough such that Assumption 3 is satisfied.

in $\mathsf{G}+\mathsf{G}^T$*. However, agent* 1 *has higher centrality than agent* 2 *in* $\mathcal{G}+\mathcal{G}^T$*.*

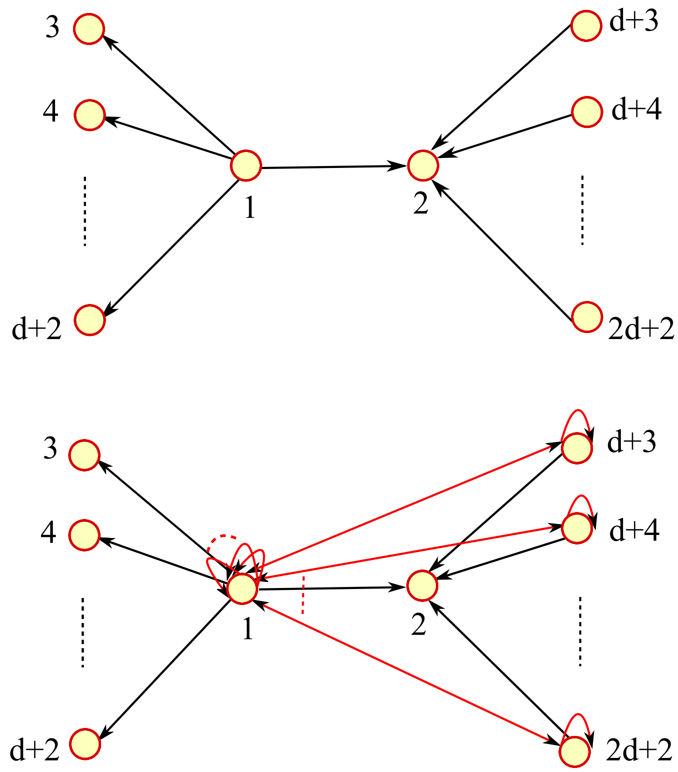

Figure 9: The top graph is G and the bottom is $\mathcal{G}$.

G Appendix: Uncertainty in the direct utility

In section 3.2 we showed when uncertainty is in the network effect (i.e., incomplete information in the strength of interactions), the maximum distortion in the whole network is due to an agent that *she and her neighbors* **together** have high Bonacich centralities in the symmetrized network $\mathsf{G}+\mathsf{G}^T$, introducing the key agents. This is a novel feature for characterizing key agents in networks. In this appendix we show that when uncertainty is in the direct utility, then the structures of the first and second-best allocations will become different than the main model where uncertainty is in the externality. Importantly, when uncertainty is in the direct utility, then the nature of consumptions and the overall distortion becomes *directly* related to Bonacich centrality in the symmetrized network, which is closely related to the previous studies (e.g., most notably Ballester et al. [2006], Candogan et al. [2012] and Bloch and Quérou [2013]).

Uncertainty in the direct utility In this case, the (ex post) utility of each agent is given by

$$u_i(\theta_i, x_i, x_{-i}, t_i) = \underbrace{\theta_i x_i - \frac{b}{2} x_i^2}_{\substack{\text{direct utility} \\ \text{(type dependent)}}} + \underbrace{x_i \sum_{j \neq i} g_{ij} x_j}_{\substack{\text{indirect utility} \\ \text{(network effect)}}} - \underbrace{t_i}_{\text{payment}}, \tag{93}$$

where x_i is the amount of the good she consumes, $x_{-i} \triangleq (x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$ is the consumption of other agents excluding agent i and t_i is the disutility charged for x_i by the firm. In addition, $\theta_i > 0$ measures the *intrinsic* marginal valuation for agent i. In the sequel, we call θ_i agent i's valuation factor, that is her type.

Assumption 4. *For each* $i = 1, 2, \cdots, n$, $b > \sum_{j \neq i}(g_{ij} + g_{ji})$.

We next characterize the first-best and the second-best solutions.

Proposition 18. *The first-best and the second-best allocations are as follows.*

(i) Under Assumption 4, the first-best trade profile is given by:

$$\mathbf{x}^{FB}(\theta) = (b\mathsf{I} - (\mathsf{G} + \mathsf{G}^T))^{-1} \boldsymbol{y}_\theta, \tag{94}$$

for any $\theta \in [\underline{\theta}\ \bar{\theta}]^n$, *where* $\mathbf{y}_\theta \triangleq diag(\theta_1, \theta_2, \cdots, \theta_n)\mathbf{1}$.

(ii) Under Assumptions 4 and 2, the second-best trade profile is given by:

$$\mathbf{x}^{SB}(\theta) = (b\mathsf{I} - (\mathsf{G} + \mathsf{G}^T))^{-1} \mathbf{y}_{\psi,\theta}, \tag{95}$$

for any $\theta \in [\underline{\theta}\ \bar{\theta}]^n$, *where* $\mathbf{y}_{\psi,\theta} \triangleq diag(\psi(\theta_1), \psi(\theta_2), \cdots, \psi(\theta_n))\mathbf{1}$.

The proof of this result follows similar steps as in the proof of Proposition 1.

In contrast to Proposition 1, where uncertainty is in the externality, here θ and $\psi(\theta)$ appear *outside* of the matrix component of the answers (compare (94), (95) with Proposition 1). This is due to the location of uncertainty (i.e. which is in the direct utility or the externality), and, thus, is critical for the following result.

Given the definitions in 2 and 3, and the result in Proposition 18 we have the following result.

Proposition 19. *For any type profile $\theta \in [\underline{\theta}\ \bar{\theta}]^n$:*

(i) Distortion vector is given by:

$$\boldsymbol{d}(\theta) = (b\mathsf{I} - (\mathsf{G} + \mathsf{G}^T))^{-1}\mathbf{y}_{\phi,\theta}, \tag{96}$$

where $\mathbf{y}_{\phi,\theta} \triangleq diag(\phi(\theta_1), \phi(\theta_2), \cdots, \phi(\theta_n))\mathbf{1}$.

(ii) Distortion is downward, *i.e., $\boldsymbol{d}(\theta) \geq \mathbf{0}$.*

(iii) Let $[\boldsymbol{d}(\theta)]_i$ denote the distortion on agent i's trade with regard to θ. Then

$$[\boldsymbol{d}(\theta)]_i > 0.$$

if there exists at least one agent j whose $\theta_j \neq \bar{\theta}$.

(iv) Let $\theta_i \leq \bar{\theta}$ and $\theta_j = \bar{\theta}$, for all $j \neq i$. Thus, $\phi(\theta_j) = 0$, for all $j \neq i$. Therefore,

$$\begin{aligned}
\mathcal{T}_i(\theta_i, \bar{\theta}_{-i}) = \mathbf{1}^T \boldsymbol{d}(\theta_i, \bar{\theta}_{-i}) &= (1/b)\left[\mathbf{1}^T\left(\mathsf{I} - \frac{1}{b}(\mathsf{G} + \mathsf{G}^T)\right)^{-1}\right]\mathbf{y}_{\phi,\theta}\mathbf{1}, \\
&= (1/b)[k_1\ k_2\ \cdots\ k_n][0\ 0\ \cdots \phi(\theta_i)\ \cdots\ 0]^T \\
&= \left(\frac{\phi(\theta_i)}{b}\right)k_i,
\end{aligned}$$

where k_i is agent i's Bonacich centrality according to $(\mathsf{I} - \frac{1}{b}(\mathsf{G} + \mathsf{G}^T))^{-1}\mathbf{1}$.

The first three parts are intuitive. However, this last part of this result shows that the maximum distortion is due to an agent who has the highest Bonacich centrality, i.e., k_i, in $\mathsf{G} + \mathsf{G}^T$ which is closely related to the previously studies (e.g., Ballester et al. [2006], Candogan et al. [2012] and Bloch and Quérou [2013]). However, in section 3.2 we showed when uncertainty is in the externality, the maximum distortion in the whole network is due to an agent that *she and her neighbors* **together** have high Bonacich centralities in the symmetrized network $\mathsf{G} + \mathsf{G}^T$, introducing the key agents. This is a novel feature for characterizing key agents in networks which is due to the uncertainty in the externality (i.e., incomplete information in the strength of interactions).

Proof of Proposition 19. The first part of the proposition directly follows from the definition 2 and Proposition 18. To prove the second and third part, define $\mathsf{M} \triangleq \left(\mathsf{I} - \frac{1}{b}(\mathsf{G} + \mathsf{G}^T)\right)^{-1}$. Thus,

$$m_{ij} = \sum_{k=0}^{\infty} \frac{1}{b^k}(g_{ij} + g_{ji})^{[k]} \tag{97}$$

where m_{ij} is the ij entry of M and $(g_{ij} + g_{ji})^{[k]}$ is the ij enry of $(\mathsf{G}_{ave})^k$. In other words, $(g_{ij} + g_{ji})^{[k]}$ is the total number of weighted walks of length k between agents i and j in G_{ave}, (see e.g., Jackson [2008]). Next, assume agent j's report is lower than the highest type, i.e., $\theta_j < \bar{\theta}$. Thus, $\phi(\theta_j) = \frac{1-F(\theta_j)}{f(\theta_j)} > 0$. Now, consider agent i, since G_{ave} is a connected network, thus agents i and j are connected through a path in G_{ave}. Further, as we shown in part (i), distortion vector is characterized as:

$$\begin{aligned} \mathbf{d}(\theta) &= (1/b)\left(\mathsf{I} - \frac{1}{b}(\mathsf{G} + \mathsf{G}^T)\right)^{-1} \ \mathbf{y}_{\phi,\theta}\mathbf{1}, \\ &= (1/b)\ \mathsf{M}\ \mathbf{y}_{\phi,\theta}\mathbf{1}, \end{aligned}$$

where $\mathbf{y}_{\phi,\theta}\mathbf{1} = (\phi(\theta_1), \phi(\theta_2), \cdots, \phi(\theta_n))^T$ (a column vector). Since i and j are connected, $m_{ij} > 0$ (due to (97)), moreover, $\phi(\theta_j) > 0$ (due to $\theta_j < \bar{\theta}$), thus:

$$[\mathbf{d}(\theta)]_i \geq m_{ij}\phi(\theta_j) > 0. \tag{98}$$

The proof of the last part is immediate by the definition. □

Finally, we wrap up this section with the following example.

Example 5 (Kite network)**.** *Consider the following symmetric network,* $g_{ij} = g_{ji} \in \{0, .1\}, i, j \in \{1, 2, \cdots, 5\}$*, that captures the interrelations among the agents. Let* $\theta_i \in \{\underline{\theta}, \bar{\theta}\}$ *with* $Prob\{\theta = \underline{\theta}\} = \frac{1}{2}$ *(b is normalized to* 1 *and* $\Delta\theta = 1$*). The following table characterizes the distortion vector when only one agent is inefficient. As shown in the table, the distortion propagates throughout the network proportional to* Bonacich centrality*. To be precise, if distortion is due to say agent* 1*, agent* 1 *exerts the highest inefficacy in its allocation, and this affects others proportional to their Bonacich centrality as in* (96) *(notice that agent* 2 *is the central agent).*

The third column of the table implies that overall distortion is proportional to Bonacich centrality of the agents in the network. Thus, distortion in agents 2's trade (the central agent) *results in the highest overall distortion in the* whole *network*[66]*. Therefore, agent 2 is the key agent.*

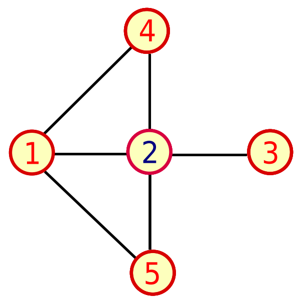

Figure 10: Kite network. Interconnection among the agents.

Inefficient agent	Distortion vector	$\mathcal{T}(\theta_i, \bar{\theta}_{-i})$
1	(**1.2** **.38** .07 .31 .31)	2.27
2	(.38 **1.25** .25 .32 .32)	**2.52**
3	(.07 **.25** **1.05** .06 .06)	1.49
4	(.31 **.32** .06 **1.12** .12)	1.93
5	(.31 **.32** .06 .12 **1.12**)	1.93

Table 1: This table shows how distortion propagates throughout the network. Inefficiency in agent 2's trade (the central agent) creates the highest distortion in the whole network.

[66]Notice that $\frac{\phi(\theta)}{b} = 1$